GenCore version 4.5

OM protein - protein search, using sw model

Run on: December 12, 2000, 01:13:49 ; Search time 30.01 Seconds
(without alignments)
5.585 Million cell updates/sec

Title: US-08-860-232-12
Perfect score: 51
Sequence: 1 VMAGVGSPYV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	51	100.0	10	1	US-08-467-083-29	Sequence 29, Appl
2	51	100.0	10	1	US-08-414-417B-29	Sequence 29, Appl
3	51	100.0	10	2	US-08-486-348A-29	Sequence 29, Appl
4	51	100.0	10	2	US-08-468-545B-29	Sequence 29, Appl
5	51	100.0	10	3	US-08-466-680B-29	Sequence 12, Appl
6	51	100.0	10	4	PCT-US95-16415-12	Sequence 263, App
7	47	92.2	10	3	US-08-159-339A-263	Sequence 44, Appl
8	38	74.5	15	1	US-08-467-083-44	Sequence 44, Appl
9	38	74.5	15	1	US-08-414-417B-44	Sequence 44, Appl
10	38	74.5	15	2	US-08-486-348A-44	Sequence 44, Appl
11	38	74.5	15	2	US-08-468-545B-44	Sequence 44, Appl
12	38	74.5	15	3	US-08-466-680B-44	Sequence 10, Appl
13	33	64.7	18	3	US-08-817-926-10	Sequence 10, Appl
14	29	56.9	9	2	US-08-725-736D-10	Sequence 10, Appl
15	29	56.9	9	3	US-09-162-368B-10	Sequence 5, Appli
16	27	52.9	10	1	US-08-172-707-5	Sequence 5, Appli
17	27	52.9	10	1	US-08-412-865-5	Sequence 5, Appli
18	27	52.9	10	1	US-08-476-505-5	Sequence 5, Appli
19	27	52.9	10	1	US-08-487-396-5	Sequence 5, Appli
20	27	52.9	10	2	US-08-941-553-5	Sequence 5, Appli
21	27	52.9	10	3	US-08-769-143-5	Sequence 1, Appli
22	27	52.9	17	4	PCT-US94-05905-1	Sequence 9, Appli
23	26	51.0	9	2	US-08-725-736D-9	Sequence 9, Appli
24	26	51.0	9	3	US-09-162-368B-9	Sequence 28, Appl
25	26	51.0	9	3	US-09-162-368B-28	Sequence 11, Appl
26	26	51.0	16	3	US-08-630-916A-11	Sequence 12, Appl
27	25	49.0	8	3	US-09-082-737-12	Sequence 8, Appli
28	25	49.0	9	2	US-08-725-736D-8	Sequence 14, Appl
29	25	49.0	9	2	US-08-725-736D-14	Sequence 36, Appl
30	25	49.0	9	2	US-08-318-856A-36	Sequence 8, Appli
31	25	49.0	9	3	US-09-162-368B-8	Sequence 14, Appl
32	25	49.0	9	3	US-09-162-368B-14	Sequence 31, Appl
33	25	49.0	9	3	US-09-162-368B-31	Sequence 6, Appli
34	25	49.0	13	1	US-08-594-447-6	Sequence 5, Appli
35	25	49.0	13	1	US-08-541-964-5	Sequence 20, Appl
36	25	49.0	13	2	US-08-665-647-20	Sequence 108, App
37	25	49.0	16	3	US-08-802-981-108	Sequence 4, Appli
38	24	47.1	9	2	US-08-725-736D-4	Sequence 7, Appli
39	24	47.1	9	2	US-08-725-736D-7	Sequence 11, Appl
40	24	47.1	9	2	US-08-725-736D-11	Sequence 8, Appli
41	24	47.1	9	3	US-08-880-963-8	Sequence 4, Appli
42	24	47.1	9	3	US-09-162-368B-4	Sequence 7, Appli
43	24	47.1	9	3	US-09-162-368B-7	Sequence 11, Appl
44	24	47.1	9	3	US-09-162-368B-11	Sequence 22, Appl
45	24	47.1	9	3	US-09-162-368B-22	Sequence 1, Appli
46	24	47.1	10	1	US-08-180-572-1	Sequence 5, Appli
47	24	47.1	10	2	US-08-725-736D-5	Sequence 6, Appli
48	24	47.1	10	2	US-08-725-736D-6	Sequence 139, App
49	24	47.1	10	2	US-08-934-222-139	Sequence 139, App
50	24	47.1	10	2	US-08-933-402-139	Sequence 139, App
51	24	47.1	10	2	US-09-207-621-139	Sequence 139, App
52	24	47.1	10	2	US-08-532-818-139	Sequence 10, Appl
53	24	47.1	10	3	US-08-880-963-10	Sequence 358, App
54	24	47.1	10	3	US-08-159-339A-358	Sequence 360, App
55	24	47.1	10	3	US-08-159-339A-360	Sequence 5, Appli
56	24	47.1	10	3	US-09-162-368B-5	Sequence 6, Appli
57	24	47.1	10	3	US-09-162-368B-6	Sequence 139, App
58	24	47.1	10	3	US-09-231-797-139	Sequence 139, App
59	24	47.1	10	3	US-08-934-224-139	Sequence 139, App
60	24	47.1	10	3	US-08-933-843-139	Sequence 10, Appl
61	24	47.1	11	1	US-08-431-539-10	Sequence 16, Appl
62	24	47.1	11	1	US-08-431-539-16	Sequence 3, Appli
63	24	47.1	11	2	US-08-725-736D-3	Sequence 3, Appli
64	24	47.1	11	3	US-09-162-368B-3	Sequence 4, Appli
65	24	47.1	12	1	US-07-958-083-4	Sequence 31, Appl
66	24	47.1	14	1	US-08-191-866D-31	Sequence 31, Appl
67	24	47.1	14	2	US-08-185-949B-31	Sequence 26, Appl
68	24	47.1	15	3	US-08-256-747C-26	Patent No. 5171839
69	24	47.1	16	5	5171839-5	Sequence 17, Appl
70	24	47.1	17	3	US-09-192-048-17	Sequence 80, Appl
71	24	47.1	19	3	US-08-802-981-80	Sequence 81, Appl
72	24	47.1	19	3	US-08-802-981-81	Sequence 5, Appli
73	24	47.1	20	1	US-07-678-974D-5	Sequence 10, Appl
74	24	47.1	20	2	US-08-945-168-10	Sequence 12, Appl
75	24	47.1	20	3	US-08-256-747C-12	

ALIGNMENTS

RESULT 1
US-08-467-083-29
; Sequence 29, Application US/08467083
; Patent No. 5726023
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/467,083
;       FILING DATE:  06-JUN-1995
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/414,417
;       FILING DATE:  06-JUN-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sharkey, Richard G.
;       REGISTRATION NUMBER:  32,629
;       REFERENCE/DOCKET NUMBER:  920010.448C2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;       TELEX:  3723836 SEEDANBERRY
;   INFORMATION FOR SEQ ID NO:  29:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  10 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
US-08-467-083-29


  Query Match             100.0%;  Score 51;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.0026;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VMAGVGSPYV 10
           ||||||||||
Db       1 VMAGVGSPYV 10


RESULT   2
US-08-414-417B-29
; Sequence 29, Application US/08414417B
; Patent No. 5801005
;  GENERAL INFORMATION:
;    APPLICANT:  Cheever, Martin A.
;    APPLICANT:  Disis, Mary L.
;    TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;    TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;    TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/414,417B
;      FILING DATE:  31-MAR-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sharkey, Richard G.
;      REGISTRATION NUMBER:  32,629
;      REFERENCE/DOCKET NUMBER:  920010.448C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-414-417B-29


  Query Match             100.0%;  Score 51;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.0026;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VMAGVGSPYV 10
           ||||||||||
Db       1 VMAGVGSPYV 10


RESULT   3
US-08-486-348A-29
; Sequence 29, Application US/08486348A
; Patent No. 5846538
;  GENERAL INFORMATION:
;    APPLICANT:  Cheever, Martin A.
;    APPLICANT:  Disis, Mary L.
;    TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;    TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;    TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,348A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sharkey, Richard G.
;      REGISTRATION NUMBER:  32,629
;      REFERENCE/DOCKET NUMBER:  920010.448C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-486-348A-29


  Query Match             100.0%;  Score 51;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.0026;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VMAGVGSPYV 10
           ||||||||||
Db       1 VMAGVGSPYV 10


RESULT   4
US-08-468-545B-29
; Sequence 29, Application US/08468545B
; Patent No. 5876712
;  GENERAL INFORMATION:
;    APPLICANT:  Cheever, Martin A.
;    APPLICANT:  Disis, Mary L.
```

```
;    TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;    TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;    TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,545B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sharkey, Richard G.
;      REGISTRATION NUMBER:  32,629
;      REFERENCE/DOCKET NUMBER:  920010.448C5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-468-545B-29


  Query Match           100.0%;  Score 51;  DB 2;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 0.0026;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VMAGVGSPYV 10
          ||||||||||
Db      1 VMAGVGSPYV 10


RESULT   5
US-08-466-680B-29
; Sequence 29, Application US/08466680B
; Patent No. 6075122
;  GENERAL INFORMATION:
;    APPLICANT:  Cheever, Martin A.
;    APPLICANT:  Disis, Mary L.
;    TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;    TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;    TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/466,680B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sharkey, Richard G.
;      REGISTRATION NUMBER:  32,629
;      REFERENCE/DOCKET NUMBER:  920010.448C4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-466-680B-29


  Query Match           100.0%;  Score 51;  DB 3;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 0.0026;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VMAGVGSPYV 10
          ||||||||||
Db      1 VMAGVGSPYV 10


RESULT   6
PCT-US95-16415-12
; Sequence 12, Application PC/TUS9516415
;  GENERAL INFORMATION:
;    APPLICANT:  The Scripps Research Institute
;    TITLE OF INVENTION:  IN VIVO ACTIVATION OF TUMOR-SPECIFIC
;    TITLE OF INVENTION:  CYTOTOXIC T CELLS
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  The Scripps Research Institute
;      STREET:  10666 North Torrey Pines Road, TPC-8
;      CITY:  La Jolla
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/16415
;      FILING DATE:  13-DEC-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/355,558
;      FILING DATE:  14-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Logan, April C.
;      REGISTRATION NUMBER:  33,950
;      REFERENCE/DOCKET NUMBER:  433.1PC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 554-2937
;      TELEFAX:  (619) 554-6312
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US95-16415-12


  Query Match           100.0%;  Score 51;  DB 4;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 0.0026;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VMAGVGSPYV 10
          ||||||||||
```

```
Db      1 VMAGVGSPYV 10


RESULT   7
US-08-159-339A-263
; Sequence 263, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  263:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-263


  Query Match             92.2%;  Score 47;  DB 3;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.013;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VMAGVGSPY 9
          |||||||||
Db      2 VMAGVGSPY 10


RESULT   8
US-08-467-083-44
; Sequence 44, Application US/08467083
; Patent No. 5726023
;  GENERAL INFORMATION:
;    APPLICANT:  Cheever, Martin A.
;    APPLICANT:  Disis, Mary L.
;    TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
;    TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;    TITLE OF INVENTION:  HER-2/NEU ONCOGENE IS ASSOCIATED
;    NUMBER OF SEQUENCES:  68
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/467,083
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/414,417
;      FILING DATE:  06-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sharkey, Richard G.
;      REGISTRATION NUMBER:  32,629
;      REFERENCE/DOCKET NUMBER:  920010.448C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;      TELEX:  3723836 SEEDANBERRY
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-467-083-44


  Query Match             74.5%;  Score 38;  DB 1;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.71;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 GVGSPYV 10
          |||||||
Db      1 GVGSPYV 7


RESULT   9
US-08-414-417B-44
; Sequence 44, Application US/08414417B
; Patent No. 5801005
;  GENERAL INFORMATION:
;    APPLICANT:  Cheever, Martin A.
;    APPLICANT:  Disis, Mary L.
;    TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;    TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;    TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/414,417B
;      FILING DATE:  31-MAR-1995
```

```
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sharkey, Richard G.
;      REGISTRATION NUMBER:  32,629
;      REFERENCE/DOCKET NUMBER:  920010.448C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-414-417B-44


  Query Match             74.5%;  Score 38;  DB 1;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.71;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 GVGSPYV 10
          |||||||
Db      1 GVGSPYV 7


RESULT  10
US-08-486-348A-44
; Sequence 44, Application US/08486348A
; Patent No. 5846538
;  GENERAL INFORMATION:
;    APPLICANT:  Cheever, Martin A.
;    APPLICANT:  Disis, Mary L.
;    TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;    TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;    TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,348A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sharkey, Richard G.
;      REGISTRATION NUMBER:  32,629
;      REFERENCE/DOCKET NUMBER:  920010.448C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-486-348A-44


  Query Match             74.5%;  Score 38;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.71;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 GVGSPYV 10
          |||||||
Db      1 GVGSPYV 7


RESULT  11
US-08-468-545B-44
; Sequence 44, Application US/08468545B
; Patent No. 5876712
;  GENERAL INFORMATION:
;    APPLICANT:  Cheever, Martin A.
;    APPLICANT:  Disis, Mary L.
;    TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;    TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;    TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,545B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sharkey, Richard G.
;      REGISTRATION NUMBER:  32,629
;      REFERENCE/DOCKET NUMBER:  920010.448C5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-468-545B-44


  Query Match             74.5%;  Score 38;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.71;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 GVGSPYV 10
          |||||||
Db      1 GVGSPYV 7


RESULT  12
US-08-466-680B-44
; Sequence 44, Application US/08466680B
; Patent No. 6075122
;  GENERAL INFORMATION:
;    APPLICANT:  Cheever, Martin A.
;    APPLICANT:  Disis, Mary L.
;    TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;    TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;    TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/466,680B
;       FILING DATE:  06-JUN-1995
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sharkey, Richard G.
;       REGISTRATION NUMBER:  32,629
;       REFERENCE/DOCKET NUMBER:  920010.448C4
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;   INFORMATION FOR SEQ ID NO:  44:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-08-466-680B-44


  Query Match             74.5%;  Score 38;  DB 3;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.71;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 GVGSPYV 10
           |||||||
Db       1 GVGSPYV 7


RESULT  13
US-08-817-926-10
; Sequence 10, Application US/08817926
; Patent No. 6001590
;  GENERAL INFORMATION:
;    APPLICANT:  Komeda, Toshihiro
;    APPLICANT:  Suda, Hisako
;    APPLICANT:  Tamai, Yukio
;    APPLICANT:  Iwamatsu, Akihiro
;    APPLICANT:  Kato, No. 6001590uo
;    APPLICANT:  Sakai, Yasuyoshi
;    TITLE OF INVENTION:  PROMOTER/TERMINATOR FOR CANDIDA BOIDINII
;    TITLE OF INVENTION:  FORMATE DEHYDROGENASE GENE
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/817,926
;      FILING DATE:  09-MAY-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP96/02597
;      FILING DATE:  12-SEP-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 234133/1995
;      FILING DATE:  12-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 42536/1996
;      FILING DATE:  29-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  081356/0112
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-817-926-10


  Query Match             64.7%;  Score 33;  DB 3;  Length 18;
  Best Local Similarity   60.0%;  Pred. No. 6.3;
  Matches    6;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 VMAGVGSPYV 10
           |:|||| ::
Db       4 VVAGVGSDHI 13


RESULT  14
US-08-725-736D-10
; Sequence 10, Application US/08725736D
; Patent No. 5831016
;  GENERAL INFORMATION:
;    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
;    TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
;    TITLE OF INVENTION:  LYMPHOCYTES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/725,736D
;      FILING DATE:  04-OCT-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/599,602
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOROTHY R. AUTH
;      REGISTRATION NUMBER:  36,434
;      REFERENCE/DOCKET NUMBER:  2026-4243
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  UNKNOWN
;    MOLECULE TYPE:
;      DESCRIPTION:  PEPTIDE
;    FEATURE:
;      NAME/KEY:  TRP-2 PEPTIDE VARIANT
;      LOCATION:
```

```
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-08-725-736D-10


  Query Match             56.9%;  Score 29;  DB 2;  Length 9;
  Best Local Similarity   62.5%;  Pred. No. 1.2e+05;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 MAGVGSPY 9
          :|| | ||
Db      1 LAGPGRPY 8


RESULT  15
US-09-162-368B-10
; Sequence 10, Application US/09162368B
; Patent No. 6083703
;  GENERAL INFORMATION:
;    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
;    TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
;    TITLE OF INVENTION:  LYMPHOCYTES
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MICROSOFT WORD 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/162,368B
;      FILING DATE:  28-SEPT-1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/725,736
;      FILING DATE:  04-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/599,602
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KATHRYN M. BROWN
;      REGISTRATION NUMBER:  34,556
;      REFERENCE/DOCKET NUMBER:  2026-4243US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  UNKNOWN
;    MOLECULE TYPE:
;      DESCRIPTION:  PEPTIDE
;    FEATURE:
;      NAME/KEY:  TRP-2 PEPTIDE VARIANT
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-09-162-368B-10


  Query Match             56.9%;  Score 29;  DB 3;  Length 9;
  Best Local Similarity   62.5%;  Pred. No. 1.2e+05;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 MAGVGSPY 9
          :|| | ||
Db      1 LAGPGRPY 8


RESULT  16
US-08-172-707-5
; Sequence 5, Application US/08172707
; Patent No. 5455168
;  GENERAL INFORMATION:
;    APPLICANT:  MARUTA, Kazuhiko
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    APPLICANT:  MIYAKE, Toshio
;    TITLE OF INVENTION:  NON-REDUCING SACCHARIDE-FORMING ENZYME,
;    TITLE OF INVENTION:  AND ITS PREPARATION AND USES
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/172,707
;      FILING DATE:  12-DEC-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 362131/1992
;      FILING DATE:  28-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 265416/1993
;      FILING DATE:  30-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  TOWNSEND, G. Kevin
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MARUTA=1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-172-707-5


  Query Match             52.9%;  Score 27;  DB 1;  Length 10;
  Best Local Similarity   83.3%;  Pred. No. 39;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 GVGSPY 9
          | ||||
Db      2 GRGSPY 7


RESULT  17
US-08-412-865-5
; Sequence 5, Application US/08412865
; Patent No. 5610047
;  GENERAL INFORMATION:
;    APPLICANT:  MARUTA, Kazuhiko
```

```
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    APPLICANT:  MIYAKE, Toshio
;    TITLE OF INVENTION:  NON-REDUCING SACCHARIDE-FORMING ENZYME,
;    TITLE OF INVENTION:  AND ITS PREPARATION AND USES
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/412,865
;      FILING DATE:  29-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/172,707
;      FILING DATE:  12-DEC-1993
;      APPLICATION NUMBER:  JP 362131/1992
;      FILING DATE:  28-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 265416/1993
;      FILING DATE:  30-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  TOWNSEND, G. Kevin
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MARUTA-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-412-865-5


  Query Match           52.9%;  Score 27;  DB 1;  Length 10;
  Best Local Similarity  83.3%;  Pred. No. 39;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 GVGSPY 9
           | ||||
Db       2 GRGSPY 7


RESULT  18
US-08-476-505-5
; Sequence 5, Application US/08476505
; Patent No. 5677442
;  GENERAL INFORMATION:
;    APPLICANT:  MARUTA, Kazuhiko
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    APPLICANT:  MIYAKE, Toshio
;    TITLE OF INVENTION:  NON-REDUCING SACCHARIDE-FORMING ENZYME,
;    TITLE OF INVENTION:  AND ITS PREPARATION AND USES
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,505
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/172,707
;      FILING DATE:  12-DEC-1993
;      APPLICATION NUMBER:  JP 362131/1992
;      FILING DATE:  28-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 265416/1993
;      FILING DATE:  30-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  TOWNSEND, G. Kevin
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MARUTA-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-476-505-5


  Query Match           52.9%;  Score 27;  DB 1;  Length 10;
  Best Local Similarity  83.3%;  Pred. No. 39;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 GVGSPY 9
           | ||||
Db       2 GRGSPY 7


RESULT  19
US-08-487-396-5
; Sequence 5, Application US/08487396
; Patent No. 5716838
;  GENERAL INFORMATION:
;    APPLICANT:  MARUTA, Kazuhiko
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    APPLICANT:  MIYAKE, Toshio
;    TITLE OF INVENTION:  NON-REDUCING SACCHARIDE-FORMING ENZYME,
;    TITLE OF INVENTION:  AND ITS PREPARATION AND USES
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,396
```

```
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/172,707
;      FILING DATE:  27-DEC-1993
;      APPLICATION NUMBER:  JP 362131/1992
;      FILING DATE:  28-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 265416/1993
;      FILING DATE:  30-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  TOWNSEND, G. Kevin
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MARUTA=1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-487-396-5


  Query Match             52.9%;  Score 27;  DB 1;  Length 10;
  Best Local Similarity   83.3%;  Pred. No. 39;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 GVGSPY 9
           | ||||
Db       2 GRGSPY 7


RESULT  20
US-08-941-553-5
; Sequence 5, Application US/08941553
; Patent No. 5922580
;  GENERAL INFORMATION:
;    APPLICANT:  MARUTA, Kazuhiko
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    APPLICANT:  MIYAKE, Toshio
;    TITLE OF INVENTION:  NON-REDUCING SACCHARIDE-FORMING ENZYME,
;    TITLE OF INVENTION:  AND ITS PREPARATION AND USES
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/941,553
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/172,707
;      FILING DATE:  12-DEC-1993
;      APPLICATION NUMBER:  JP 362131/1992
;      FILING DATE:  28-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 265416/1993
;      FILING DATE:  30-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  TOWNSEND, G. Kevin
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MARUTA=1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-941-553-5


  Query Match             52.9%;  Score 27;  DB 2;  Length 10;
  Best Local Similarity   83.3%;  Pred. No. 39;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 GVGSPY 9
           | ||||
Db       2 GRGSPY 7


RESULT  21
US-08-769-143-5
; Sequence 5, Application US/08769143
; Patent No. 6017899
;  GENERAL INFORMATION:
;    APPLICANT:  MARUTA, Kazuhiko
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    APPLICANT:  MIYAKE, Toshio
;    TITLE OF INVENTION:  NON-REDUCING SACCHARIDE-FORMING ENZYME,
;    TITLE OF INVENTION:  AND ITS PREPARATION AND USES
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/769,143
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/172,707
;      FILING DATE:  12-DEC-1993
;      APPLICATION NUMBER:  JP 362131/1992
;      FILING DATE:  28-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 265416/1993
;      FILING DATE:  30-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  TOWNSEND, G. Kevin
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MARUTA=1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
```

```
;       LENGTH:  10 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-769-143-5


  Query Match             52.9%;  Score 27;  DB 3;  Length 10;
  Best Local Similarity   83.3%;  Pred. No. 39;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 GVGSPY 9
           | ||||
Db       2 GRGSPY 7


RESULT  22
PCT-US94-05905-1
; Sequence 1, Application PC/TUS9405905
;  GENERAL INFORMATION:
;     APPLICANT:
;     TITLE OF INVENTION:  tRNA BINDING-DEPENDENT INHIBITION OF MICROBIAL
;     TITLE OF INVENTION:  PATHOGEN GROWTH
;     NUMBER OF SEQUENCES:  22
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;       STREET:  Two Militia Drive
;       CITY:  Lexington
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02173-4799
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US94/05905
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/068,382
;       FILING DATE:  28-MAY-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Brook, David E.
;       REGISTRATION NUMBER:  22,592
;       REFERENCE/DOCKET NUMBER:  MIT-6299A PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-861-6240
;       TELEFAX:  617-861-9540
;       TELEX:  951794
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  17 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
PCT-US94-05905-1


  Query Match             52.9%;  Score 27;  DB 4;  Length 17;
  Best Local Similarity   62.5%;  Pred. No. 66;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 AGVGSPYV 10
           :|  ||||
Db       1 SGYDSPYV 8


RESULT  23
US-08-725-736D-9
; Sequence 9, Application US/08725736D
; Patent No. 5831016
;  GENERAL INFORMATION:
;     APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;     TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
;     TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
;     TITLE OF INVENTION:  LYMPHOCYTES
;     NUMBER OF SEQUENCES:  20
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;       STREET:  345 PARK AVENUE
;       CITY:  NEW YORK
;       STATE:  NEW YORK
;       COUNTRY:  USA
;       ZIP:  10154
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  FLOPPY DISK
;       COMPUTER:  IBM PC COMPATIBLE
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  WORDPERFECT 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/725,736D
;       FILING DATE:  04-OCT-1996
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/599,602
;       FILING DATE:  09-FEB-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  DOROTHY R. AUTH
;       REGISTRATION NUMBER:  36,434
;       REFERENCE/DOCKET NUMBER:  2026-4243
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 758-4800
;       TELEFAX:  (212) 751-6849
;       TELEX:  421792
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9
;       TYPE:  AMINO ACID
;       STRANDEDNESS:  SINGLE
;       TOPOLOGY:  UNKNOWN
;     MOLECULE TYPE:
;       DESCRIPTION:  PEPTIDE
;     FEATURE:
;       NAME/KEY:  TRP-2 PEPTIDE VARIANT
;       LOCATION:
;       IDENTIFICATION METHOD:
;       OTHER INFORMATION:
US-08-725-736D-9


  Query Match             51.0%;  Score 26;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 MAGVGSPY 9
           ::| | ||
Db       1 LSGPGRPY 8


RESULT  24
US-09-162-368B-9
; Sequence 9, Application US/09162368B
; Patent No. 6083703
;  GENERAL INFORMATION:
;     APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;     TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
;     TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
;     TITLE OF INVENTION:  LYMPHOCYTES
;     NUMBER OF SEQUENCES:  31
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;       STREET:  345 PARK AVENUE
;       CITY:  NEW YORK
;       STATE:  NEW YORK
```

```
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MICROSOFT WORD 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/162,368B
;      FILING DATE:  28-SEPT-1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/725,736
;      FILING DATE:  04-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/599,602
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KATHRYN M. BROWN
;      REGISTRATION NUMBER:  34,556
;      REFERENCE/DOCKET NUMBER:  2026-4243US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  UNKNOWN
;    MOLECULE TYPE:
;      DESCRIPTION:  PEPTIDE
;    FEATURE:
;      NAME/KEY:  TRP-2 PEPTIDE VARIANT
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-09-162-368B-9

  Query Match           51.0%;  Score 26;  DB 3;  Length 9;
  Best Local Similarity 50.0%;  Pred. No. 1.2e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 MAGVGSPY 9
            ::| | ||
Db       1 LSGPGRPY 8


RESULT  25
US-09-162-368B-28
; Sequence 28, Application US/09162368B
; Patent No. 6083703
  GENERAL INFORMATION:
    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
    TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
    TITLE OF INVENTION:  LYMPHOCYTES
    NUMBER OF SEQUENCES:  31
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
      STREET:  345 PARK AVENUE
      CITY:  NEW YORK
      STATE:  NEW YORK
      COUNTRY:  USA
      ZIP:  10154
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  FLOPPY DISK
      COMPUTER:  IBM PC COMPATIBLE
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  MICROSOFT WORD 97
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/09/162,368B
      FILING DATE:  28-SEPT-1998
      CLASSIFICATION:  530
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  08/725,736
      FILING DATE:  04-OCT-1996
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  08/599,602
      FILING DATE:  09-FEB-1996
    ATTORNEY/AGENT INFORMATION:
      NAME:  KATHRYN M. BROWN
      REGISTRATION NUMBER:  34,556
      REFERENCE/DOCKET NUMBER:  2026-4243US1
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (212) 758-4800
      TELEFAX:  (212) 751-6849
      TELEX:  421792
  INFORMATION FOR SEQ ID NO:  28:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  9
      TYPE:  AMINO ACID
      STRANDEDNESS:  UNKNOWN
      TOPOLOGY:  UNKNOWN
    MOLECULE TYPE:
      DESCRIPTION:  PEPTIDE
    FEATURE:
      NAME/KEY:
      LOCATION:
      IDENTIFICATION METHOD:
      OTHER INFORMATION:
US-09-162-368B-28

  Query Match           51.0%;  Score 26;  DB 3;  Length 9;
  Best Local Similarity 50.0%;  Pred. No. 1.2e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 MAGVGSPY 9
            : | |:||
Db       1 LLGPGAPY 8


RESULT  26
US-08-630-916A-11
; Sequence 11, Application US/08630916A
; Patent No. 6011137
  GENERAL INFORMATION:
    APPLICANT:  Pirozzi, Gregorio
    APPLICANT:  Kay, Brian K.
    APPLICANT:  Fowlkes, Dana M.
    TITLE OF INVENTION:  IDENTIFICATION AND ISOLATION OF NOVEL
    TITLE OF INVENTION:  POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
    NUMBER OF SEQUENCES:  124
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Pennie & Edmonds
      STREET:  1155 Avenue of the Americas
      CITY:  New York
      STATE:  New York
      COUNTRY:  United States
      ZIP:  10036-2711
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/630,916A
      FILING DATE:  03-APR-1996
      CLASSIFICATION:  435
    ATTORNEY/AGENT INFORMATION:
      NAME:  MISROCK, S. LESLIE
      REGISTRATION NUMBER:  18,872
      REFERENCE/DOCKET NUMBER:  1101-203
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 896-8864/9741
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-630-916A-11


  Query Match             51.0%;  Score 26;  DB 3;  Length 16;
  Best Local Similarity   57.1%;  Pred. No. 93;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 AGVGSPY 9
           :| |:||
Db       3 SGPGTPY 9


RESULT  27
US-09-082-737-12
; Sequence 12, Application US/09082737
; Patent No. 6013500
;  GENERAL INFORMATION:
;    APPLICANT:  Minden, Audrey
;    TITLE OF INVENTION:  PAK4; A No. 6013500el Gene Encoding A Serine/
;    TITLE OF INVENTION:  Threonine Kinase
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11230
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/082,737
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0575/55311
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 391-0525
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-082-737-12


  Query Match             49.0%;  Score 25;  DB 3;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 1.2e+05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 VGSPY 9
           ||:||
Db       1 VGTPY 5


RESULT  28
US-08-725-736D-8
; Sequence 8, Application US/08725736D
; Patent No. 5831016
;  GENERAL INFORMATION:
;    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
;    TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
;    TITLE OF INVENTION:  LYMPHOCYTES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/725,736D
;      FILING DATE:  04-OCT-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/599,602
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOROTHY R. AUTH
;      REGISTRATION NUMBER:  36,434
;      REFERENCE/DOCKET NUMBER:  2026-4243
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  UNKNOWN
;    MOLECULE TYPE:
;      DESCRIPTION:  PEPTIDE
;    FEATURE:
;      NAME/KEY:  TRP-2 PEPTIDE VARIANT
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-08-725-736D-8


  Query Match             49.0%;  Score 25;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+05;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MAGVGSPY 9
           : | | ||
Db       1 LVGPGRPY 8


RESULT  29
US-08-725-736D-14
; Sequence 14, Application US/08725736D
; Patent No. 5831016
;  GENERAL INFORMATION:
;    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
;    TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
;    TITLE OF INVENTION:  LYMPHOCYTES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;       STREET:  345 PARK AVENUE
;       CITY:  NEW YORK
;       STATE:  NEW YORK
;       COUNTRY:  USA
;       ZIP:  10154
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  FLOPPY DISK
;       COMPUTER:  IBM PC COMPATIBLE
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  WORDPERFECT 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/725,736D
;       FILING DATE:  04-OCT-1996
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/599,602
;       FILING DATE:  09-FEB-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  DOROTHY R. AUTH
;       REGISTRATION NUMBER:  36,434
;       REFERENCE/DOCKET NUMBER:  2026-4243
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 758-4800
;       TELEFAX:  (212) 751-6849
;       TELEX:  421792
;   INFORMATION FOR SEQ ID NO:  14:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9
;       TYPE:  AMINO ACID
;       STRANDEDNESS:  SINGLE
;       TOPOLOGY:  UNKNOWN
;     MOLECULE TYPE:
;       DESCRIPTION:  PEPTIDE
;     FEATURE:
;       NAME/KEY:  TRP-2 PEPTIDE VARIANT
;       LOCATION:
;       IDENTIFICATION METHOD:
;       OTHER INFORMATION:
US-08-725-736D-14


  Query Match             49.0%;  Score 25;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+05;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MAGVGSPY 9
            : | | ||
Db       1 LLGPGKPY 8


RESULT  30
US-08-318-856A-36
; Sequence 36, Application US/08318856A
; Patent No. 5972351
;  GENERAL INFORMATION:
;     APPLICANT:  Adrian V.S. Hill, et al.
;     TITLE OF INVENTION:  PLASMODIUM FALCIPARUM MHC CLASS I-
;     TITLE OF INVENTION:  RESTRICTED CTL EPITOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE
;     TITLE OF INVENTION:  ANTIGENS (AS AMENDED)
;     NUMBER OF SEQUENCES:  86
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Wenderoth, Lind & Ponack, L.L.P.
;       STREET:  2033 K Street, N.W., Suite 800
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  U.S.A.
;       ZIP:  20006
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk, 3.5 inch, 1.44 mb
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  WordPerfect 5.1+
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/318,856A
;       FILING DATE:  October 3, 1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  GB 92 08 068.8
;       FILING DATE:  April 3, 1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  GB 92 17 704.7
;       FILING DATE:  August 20, 1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/GB93/00711
;       FILING DATE:  April 5, 1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Lee Cheng
;       REGISTRATION NUMBER:  40,949
;       REFERENCE/DOCKET NUMBER:  263-PPIR1577US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 721-8200
;       TELEFAX:  (202) 721-8250
;   INFORMATION FOR SEQ ID NO:  36:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acid residues
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-318-856A-36


  Query Match             49.0%;  Score 25;  DB 2;  Length 9;
  Best Local Similarity   83.3%;  Pred. No. 1.2e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 AGVGSP 8
            || |||
Db       3 AGKGSP 8


RESULT  31
US-09-162-368B-8
; Sequence 8, Application US/09162368B
; Patent No. 6083703
;  GENERAL INFORMATION:
;     APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;     TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
;     TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
;     TITLE OF INVENTION:  LYMPHOCYTES
;     NUMBER OF SEQUENCES:  31
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;       STREET:  345 PARK AVENUE
;       CITY:  NEW YORK
;       STATE:  NEW YORK
;       COUNTRY:  USA
;       ZIP:  10154
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  FLOPPY DISK
;       COMPUTER:  IBM PC COMPATIBLE
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  MICROSOFT WORD 97
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/162,368B
;       FILING DATE:  28-SEPT-1998
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/725,736
;       FILING DATE:  04-OCT-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/599,602
;       FILING DATE:  09-FEB-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  KATHRYN M. BROWN
;       REGISTRATION NUMBER:  34,556
```

```
;       REFERENCE/DOCKET NUMBER:  2026-4243US1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 758-4800
;       TELEFAX:  (212) 751-6849
;       TELEX:  421792
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9
;       TYPE:  AMINO ACID
;       STRANDEDNESS:  SINGLE
;       TOPOLOGY:  UNKNOWN
;     MOLECULE TYPE:
;       DESCRIPTION:  PEPTIDE
;     FEATURE:
;       NAME/KEY:  TRP-2 PEPTIDE VARIANT
;       LOCATION:
;       IDENTIFICATION METHOD:
;       OTHER INFORMATION:
US-09-162-368B-8


  Query Match             49.0%;  Score 25;  DB 3;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+05;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MAGVGSPY 9
           : | | ||
Db       1 LVGPGRPY 8


RESULT  32
US-09-162-368B-14
; Sequence 14, Application US/09162368B
; Patent No. 6083703
;  GENERAL INFORMATION:
;    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
;    TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
;    TITLE OF INVENTION:  LYMPHOCYTES
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MICROSOFT WORD 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/162,368B
;      FILING DATE:  28-SEPT-1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/725,736
;      FILING DATE:  04-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/599,602
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KATHRYN M. BROWN
;      REGISTRATION NUMBER:  34,556
;      REFERENCE/DOCKET NUMBER:  2026-4243US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  UNKNOWN
;    MOLECULE TYPE:
;      DESCRIPTION:  PEPTIDE
;    FEATURE:
;      NAME/KEY:  TRP-2 PEPTIDE VARIANT
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-09-162-368B-14


  Query Match             49.0%;  Score 25;  DB 3;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+05;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MAGVGSPY 9
           : | | ||
Db       1 LLGPGKPY 8


RESULT  33
US-09-162-368B-31
; Sequence 31, Application US/09162368B
; Patent No. 6083703
;  GENERAL INFORMATION:
;    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
;    TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
;    TITLE OF INVENTION:  LYMPHOCYTES
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MICROSOFT WORD 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/162,368B
;      FILING DATE:  28-SEPT-1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/725,736
;      FILING DATE:  04-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/599,602
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KATHRYN M. BROWN
;      REGISTRATION NUMBER:  34,556
;      REFERENCE/DOCKET NUMBER:  2026-4243US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  UNKNOWN
;      TOPOLOGY:  UNKNOWN
;    MOLECULE TYPE:
;      DESCRIPTION:  PEPTIDE
;    FEATURE:
;      NAME/KEY:  TRP-2 peptide Variant
;      LOCATION:  2 to 9
```

```
;      IDENTIFICATION METHOD:  Experimental
;      OTHER INFORMATION:  First Xaa is Ile, Ser,
;      OTHER INFORMATION:  Leu, Val; Second Xaa is Lys or Arg
US-09-162-368B-31


  Query Match             49.0%;  Score 25;  DB 3;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+05;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MAGVGSPY 9
          : | | ||
Db      1 LXGPGRPY 8


RESULT  34
US-08-594-447-6
; Sequence 6, Application US/08594447
; Patent No. 5776716
;  GENERAL INFORMATION:
;    APPLICANT:  Ron, Dorit
;    APPLICANT:  Napolitano, Eugene W.
;    APPLICANT:  Voronova, Anna F.
;    TITLE OF INVENTION:  METHODS FOR IDENTIFYING AGENTS WHICH
;    TITLE OF INVENTION:  BLOCKTHE INTERACTION OF FYN WITH PKC-THETA, AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  75
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW - Ste. 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/594,447
;      FILING DATE:  31-JAN-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  22550-20025.24
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 822-0168
;      TELEX:  90-4030 MRSNFOERSWSH
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..13
;      OTHER INFORMATION:  /label= beta-C2-2
US-08-594-447-6


  Query Match             49.0%;  Score 25;  DB 1;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GVGSPYV 10
          |:  |||
Db      5 GLSDPYV 11


RESULT  35
US-08-541-964-5
; Sequence 5, Application US/08541964
; Patent No. 5783405
;  GENERAL INFORMATION:
;    APPLICANT:  Mochly-Rosen, Daria
;    APPLICANT:  Ron, Dorit
;    APPLICANT:  Kauvar, Lawrence M.
;    APPLICANT:  Napolitano, Eugene W.
;    TITLE OF INVENTION:  A RAPID SCREENING METHOD FOR EFFECTORS
;    TITLE OF INVENTION:  OF SIGNAL TRANSDUCTION
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 PENNSYLVANIA AVENUE, NW-STE. 5500
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/541,964
;      FILING DATE:  10-OCT-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  22550-20025.23
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 822-0168
;      TELEX:  90-4030 MRSNFOERSWSH
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..13
;      OTHER INFORMATION:  /label= beta-C2-2
US-08-541-964-5


  Query Match             49.0%;  Score 25;  DB 1;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GVGSPYV 10
          |:  |||
Db      5 GLSDPYV 11


Search completed: December 12, 2000, 02:44:01
Job time: 5412 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: December 11, 2000, 23:07:30 ; Search time 47.49 Seconds
(without alignments)
7.200 Million cell updates/sec

Title: US-08-860-232-12
Perfect score: 51
Sequence: 1 VMAGVGSPYV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 127583

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : A_Geneseq_36:*
1: /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	51	100.0	10	15	R61525	Peptide fragment (
2	51	100.0	10	17	R97508	Cytotoxic T lympho
3	51	100.0	10	18	W36826	Immunogenic peptid
4	51	100.0	10	19	W77132	HER-2/neu syntheti
5	51	100.0	10	19	W70071	HER-2/neu derived
6	51	100.0	15	21	Y98861	HLA class II bindi
7	47	92.2	9	20	Y46001	Immunogenic peptid
8	47	92.2	9	20	Y46413	Immunogenic peptid
9	47	92.2	9	20	Y46478	Immunogenic peptid
10	47	92.2	10	15	Y37958	Human cERB2 oncoge
11	47	92.2	10	20	Y45527	Immunogenic peptid
12	33	64.7	15	21	Y98860	HLA class II bindi
13	29	56.9	9	18	W30834	TRP-2 derived pote
14	29	56.9	15	21	Y98957	HLA class II bindi
15	26	51.0	9	18	W30833	TRP-2 derived pote
16	26	51.0	9	18	W37019	TRP-2 derived pote
17	26	51.0	16	18	W38105	Peptide recognitio.
18	25	49.0	6	17	Y07541	Hexapeptide having
19	25	49.0	6	19	W11180	Cyclic antiarrhyth
20	25	49.0	8	17	R95514	Human prostate car
21	25	49.0	8	21	Y59138	Human PAK65 kinase
22	25	49.0	9	18	W30838	TRP-2 derived pote
23	25	49.0	9	18	W30832	TRP-2 derived pote
24	25	49.0	12	15	R57426	Rabphilin-3A fragm
25	25	49.0	12	19	W62130	Haemophilus influe
26	25	49.0	13	18	W15779	Protein kinase C-b
27	25	49.0	13	21	Y66838	T cell antigen rec
28	25	49.0	15	19	W85388	Helper T-cell clas
29	25	49.0	15	19	W85376	Helper T-cell clas
30	25	49.0	15	19	W85361	Helper T-cell clas
31	25	49.0	15	19	W85181	Helper T-cell pept
32	25	49.0	15	19	W85164	Helper T-cell pept
33	25	49.0	15	19	W85155	Helper T-cell pept
34	25	49.0	15	20	Y55981	Human SULU1-derive
35	25	49.0	15	21	Y88496	Peptide #2 used in
36	25	49.0	15	21	Y88500	Peptide #6 used in
37	25	49.0	15	21	Y88520	Peptide #26 used i
38	25	49.0	15	21	Y73273	Plasmodium falcipa
39	25	49.0	16	19	W82182	Fluorogenic protea
40	25	49.0	16	20	Y55977	Human PAK5-derived
41	25	49.0	17	21	Y82607	Human p21-protein
42	25	49.0	17	21	Y78109	Human p65 peptide.
43	25	49.0	18	19	W44900	Bovine "polyprolin
44	24	47.1	5	20	Y24248	Peptide #2. Synth
45	24	47.1	7	20	W74251	HJ loop peptide HJ
46	24	47.1	8	20	Y47034	Immunogenic peptid
47	24	47.1	9	18	W30828	TRP-2 derived pote
48	24	47.1	9	18	W30831	TRP-2 derived pote
49	24	47.1	9	18	W30835	TRP-2 derived pote
50	24	47.1	9	18	W37013	TRP-2 derived pote
51	24	47.1	9	20	Y53536	Human melanoma TRP
52	24	47.1	9	20	Y46785	Immunogenic peptid
53	24	47.1	9	20	Y46792	Immunogenic peptid
54	24	47.1	9	20	Y40223	Amino acid sequenc
55	24	47.1	9	20	Y26879	Melanoma-derived l
56	24	47.1	9	20	W98933	Human leukocyte an
57	24	47.1	9	21	Y54234	HLA binding peptid
58	24	47.1	9	21	Y54235	HLA binding peptid
59	24	47.1	9	21	Y73066	Hepatitis B virus
60	24	47.1	10	14	R41734	Antlerin OT-1 N-te
61	24	47.1	10	15	Y38057	Hepatitis B virus-
62	24	47.1	10	15	Y38059	Hepatitis B virus-
63	24	47.1	10	15	R67146	Calcitonin derived
64	24	47.1	10	16	R73002	O12 antigen proteo
65	24	47.1	10	18	W30829	TRP-2 derived pote
66	24	47.1	10	18	W30830	TRP-2 derived pote
67	24	47.1	10	20	Y45626	Immunogenic peptid
68	24	47.1	10	20	Y45628	Immunogenic peptid
69	24	47.1	10	20	W98926	Human leukocyte an
70	24	47.1	11	18	W30827	TRP-2 derived pote
71	24	47.1	13	15	R62441	Peptide fragment o
72	24	47.1	13	19	W62289	Glutathionylspermi
73	24	47.1	15	14	R39850	E1 peptide RV-EP26
74	24	47.1	15	18	W42275	Biotinylated cross
75	24	47.1	15	20	Y55979	Human KHS1-derived

ALIGNMENTS

RESULT 1
R61525
ID R61525 standard; peptide; 10 AA.
XX
AC R61525;

XX
DT 11-MAY-1995 (first entry)
XX
DE Peptide fragment (1.0738) of c-ERB2 binds HLA-A2.1.
XX
KW antigen; epitope; immunogenic target protein; PSA; HBVc; HBVs; EBV;
KW HIV1; plasma specific antigen; hepatitis B virus; Epstein Barr;
KW human immunodeficiency virus; human papilloma virus; p53; c-ERB2;
KW MAGE-1; melanoma antigen-1; core antigen; surface antigen;
KW pharmaceutical composition; in vivo; ex vivo; therapeutic;
KW diagnostic; MHC class I molecule; major histocompatability complex;
KW HLA-A2.1; 9mer; 10mer; anchor; human leukocyte antigen.
XX
OS Homo sapiens.
XX
PN WO9420127-A.
XX
PD 15-SEP-1994.
XX
PF 04-MAR-1994; 94WO-US02353.
XX
PR 05-MAR-1993; 93US-0027146.
PR 04-JUN-1993; 93US-0073205.
PR 29-NOV-1993; 93US-0159184.
XX
PA (CYTE-) CYTEL CORP.
XX
PI Grey HM, Kast WM, Sette A, Sidney J;
XX
DR WPI; 1994-302678/37.
XX
PT Immunogenic peptide(s) having an HLA-A2.1 binding motif - used
PT for treatment or prophylaxis of cancer, virus infection or
PT autoimmune diseases.
XX
PS Example 5; Page 108; 138pp; English.
XX
CC R59496-R61666 are immunogenic 10mer peptides that contain a HLA-A2.1
CC binding motif. These peptides bind HLA-A2.1 and have a binding
CC affinity of at least 1% as compared to a reference peptide (R71293).
CC R61525 has an IC50 of 0.018 and the sequence occurs at position 773
CC in the human c-ERB2 gene product. Peptides of the invention can
CC induce cytotoxic T lymphocytes which can react with target cells.
CC They can be used for the treatment or prophlyaxis of cancer, eg.
CC prostate cancer or lymphoma, etc.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 51; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VMAGVGSPYV 10
||||||||||
Db 1 vmagvgspyv 10

RESULT 2
R97508
ID R97508 standard; peptide; 10 AA.
XX
AC R97508;
XX
DT 11-FEB-1997 (first entry)
XX
DE Cytotoxic T lymphocyte-activating Her-2/Neu-specific peptide.
XX
KW p53; Her-2; Neu; aa; amino acid; CTL; cytotoxic T lymphocyte; target;
KW malignant cell; antigenic; vaccine; immunisation; activation.
XX
OS Homo sapiens.
XX
PN WO9618409-A1.
XX
PD 20-JUN-1996.
XX
PF 14-DEC-1995; 95WO-US16415.
XX
PR 14-DEC-1994; 94US-0355558.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Sherman LA;
XX
DR WPI; 1996-300385/30.
XX
PT In vivo activation of tumour-specific cytotoxic T lymphocytes - by
PT contacting with polypeptide(s) derived from human p53 or Her-2/Neu
PT proteins
XX
PS Claim 5; Page 124; 158pp; English.
XX
CC R97508 is a peptide capable of activating cytotoxic T lymphocytes
CC (CTLs) which specifically target malignant cells. The peptide
CC corresponds to amino acids 773-782 of human Her-2/Neu protein. CTL-
CC activating peptides can be used in a vaccine for protecting against
CC tumour cell formation. CTLs activated by the peptides will lyse
CC tumour cells displaying specific peptides. Antibodies against CTL-
CC activating peptides are useful for the identification of other
CC similar compounds which may be useful for treating cancer or virally-
CC infected cells, or for diagnosis. The peptide and vaccines produced
CC provide immunity to a high percentage of different ethnic groups,
CC i.e. those with different HLA alleles.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 51; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VMAGVGSPYV 10
||||||||||
Db 1 vmagvgspyv 10

RESULT 3
W36826
ID W36826 standard; peptide; 10 AA.
XX
AC W36826;
XX
DT 23-MAR-1998 (first entry)
XX
DE Immunogenic peptide H7 based on the human Her-2/neu protein.
XX
KW Her-2/neu protein; human leukocyte antigen A2.1; HLA;
KW cytotoxic T lymphocyte; CTL; immune response; tumour-associated antigen;
KW T-cell receptor; TCR; tumour treatment.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9732603-A1.
XX
PD 12-SEP-1997.
XX
PF 05-MAR-1997; 97WO-US03611.
XX
PR 05-MAR-1996; 96US-0012845.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Lustgarten J, Sherman LA;
XX

DR WPI; 1997-470496/43.
XX
PT Nucleic acid encoding variable regions of HLA-restricted non-human T
PT cell receptor specific for tumour antigen - used to identify tumour
PT antigens and for tumour therapy
XX
PS Example 1; Page 9; 34pp; English.
XX
CC Synthetic peptides W36824-40 are based on the sequence of the human
CC Her-2/neu protein, wherein each sequence contains the anchor motif for
CC human leukocyte antigen (HLA) A2.1. The present peptide is based on
CC positions 773-782. The ability of these peptides to inhibit the binding
CC of an influenza virus matrix protein peptide M1 to HLA A2.1 was measured
CC by inhibition of lysis by an M1 specific, HLA A2.1 restricted, cytotoxic
CC T lymphocyte (CTL) clone. The present protein showed 55% inhibition. The
CC peptides were also tested for their ability to elicit an immune response
CC in vivo. However, only H3 (W36824) and H7 (W36826) were able to do
CC this. H3 and H7 peptides are tumour-associated antigens, and were used to
CC immunize a transgenic, non-human vertebrate (that has been modified to
CC express at least one HLA antigen), so that the animal produces CTL which
CC displays HLA-restricted T-cell receptor (TCR) specificity for the
CC antigen. Nucleic acid encoding variable regions of the alpha and beta
CC chains of such TCRs can be amplified from CTLs produced in the above
CC manner. Cells expressing recombinant TCR are used to identify antigens
CC associated with a tumour and to treat tumours in humans. Transgenic mice
CC are a more convenient source of CTL than the tumour-infiltrating
CC lymphocytes previously used. TCR can be humanised to reduce
CC side-reactions and short peptide derivatives of TCR are more economical
CC to produce than TCR itself, particularly when expressed as a
CC single-chain molecule rather than as a dimer.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 51; DB 18; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VMAGVGSPYV 10
||||||||||
Db 1 vmagvgspyv 10

RESULT 4
W77132
ID W77132 standard; peptide; 10 AA.
XX
AC W77132;
XX
DT 16-NOV-1998 (first entry)
XX
DE HER-2/neu synthetic peptide epitope 2.
XX
KW Tyrosinase; tyrosinase cytotoxic lymphocyte response;
KW cytotoxic T lymphocyte; cysteine-depleted; melanoma.
XX
OS Synthetic.
XX
PN WO9833810-A2.
XX
PD 06-AUG-1998.
XX
PF 29-JAN-1998; 98WO-US01592.
XX
PR 30-JAN-1997; 97US-0037781.
XX
PA (UYVI-) UNIV VIRGINIA PATENT FOUND.
XX
PI Engelhard VH, Hunt DF, Kittlesen D, Slingluff CL;
XX
DR WPI; 1998-437388/37.
XX
PT Disease specific immunogen - comprises disease specific cytotoxic T
PT lymphocyte epitope used to elicit melanoma specific CTL response
XX
PS Disclosure; Page 27; 93pp; English.
XX
CC The peptide epitope W77119-W77138 were created for human tumour-specific
CC cytotoxic T lymphocyte response. These peptides are are cysteine-
CC depleted mutants of a native disease-specific CTL epitope. The cysteine-
CC depleted CTL epitopes elicit a stronger or more specific CTL response
CC than the native epitope. The epitopes can be used in a disease-specific
CC immunogen to protect a mammal against disease in particular melanomas.
CC The peptides may also be used to screen a sample for the presence of
CC an antigen with the same epitope, or with a different cross-reactive
CC epitope.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 51; DB 19; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VMAGVGSPYV 10
||||||||||
Db 1 vmagvgspyv 10

RESULT 5
W70071
ID W70071 standard; peptide; 10 AA.
XX
AC W70071;
XX
DT 22-OCT-1998 (first entry)
XX
DE HER-2/neu derived HLA-A2.1 binding peptide 19 (residues 773-782).
XX
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW human leukocyte antigen; HLA; tumour associated antigen; cancer;
KW antigen presenting cell; APC; immunogenic peptide; immune disorder;
KW viral infection; AIDS; hepatitis; bacterial infection; malaria;
KW fungal infection; tuberculosis; melanoma; HER-2/neu; cerB-2.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9833888-A1.
XX
PD 06-AUG-1998.
XX
PF 30-JAN-1998; 98WO-US01959.
XX
PR 31-JAN-1997; 97US-0036696.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Celis E, Sette A, Sidney J, Southwood S, Tsai V;
XX
DR WPI; 1998-437445/37.
XX
PT Production of antigen-specific cytotoxic T cells - by incubating
PT immunogenic peptide(s) from antigen that binds class I major
PT histocompatibility complex molecules with pre-treated antigen
PT presenting cells
XX
PS Example 7; Page 77; 104pp; English.
XX
CC Sequences shown in W70053 to W70075 represent peptides derived from
CC HER-2/neu (cerB-2) antigen. The peptides can bind to a human leukocyte
CC antigen (HLA), HLA-A2.1 and are used to exemplify the method of
CC invention of producing antigen-specific cytotoxic T cells (CTLs) in
CC vitro. The method comprises contacting immunogenic peptides from an
CC antigen that binds class I major histocompatibility complex (MHC)
CC molecules with antigen presenting cells (APCs) pretreated with

CC pretreatment growth factors, and incubating the APCs with purified CD8
CC cells in the presence of at least 2 incubation growth factors, thereby
CC producing antigen-specific CTLs. A method for specifically killing
CC target cells in a human patient is also provided which comprises
CC obtaining a fluid sample containing CTLs from a patient, contacting the
CC cytotoxic T cells with APCs pretreated with pre-treatment growth factors,
CC where the APCs comprise class I MHC molecules. The pretreated APCs are
CC incubated with the cytotoxic growth factors, thereby producing activated
CC CTLs which are contacted with a carrier to form a composition. The
CC composition can then be administered to the patient. The activated CTLs
CC can be used for treating cancers, immune disorders, viral infections,
CC AIDS, hepatitis, bacterial infection, fungal infection, malaria or
CC tuberculosis.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 51; DB 19; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VMAGVGSPYV 10
||||||||||
Db 1 vmagvgspyv 10

RESULT 6
Y98861
ID Y98861 standard; Peptide; 15 AA.
XX
AC Y98861;
XX
DT 07-AUG-2000 (first entry)
XX
DE HLA class II binding antigen epitope peptide #50.
XX
KW Human leucocyte antigen; HLA class II; antigen epitope; pharmaceutical;
KW immune response; chronic viral disease; cancer; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; myasthenia gravis; AIDS;
KW allograft rejection; allergy; lyme disease; hepatitis; prostate cancer;
KW glomerulonephritis; food hypersensitivity; malaria.
XX
OS Unidentified.
XX
PN WO9961916-A1.
XX
PD 02-DEC-1999.
XX
PF 28-MAY-1999; 99WO-US12066.
XX
PR 29-MAY-1998; 98US-0087192.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Southwood S, Sidney J;
XX
DR WPI; 2000-097143/08.
XX
PT New compositions containing immunogenic peptide epitopes for various
PT HLA class II DR molecules useful for inducing helper T cell response -
XX
PS Claim 1; Page 40; 60pp; English.
XX
CC The present invention relates to a new pharmaceutical composition
CC comprising a unit dose form of a peptide, or analogue, comprising an
CC epitope selected from those represented by peptides Y98812-Y99339 which
CC are derived from various antigens for various human leucocyte antigen
CC class DR molecules, representative of the world wide population. The
CC peptide/analogue binds to an HLA class II molecule at an IC-50 of less
CC than or equal to 1,000 nM. The pharmaceutical can be used to induce a
CC helper T cell response. The pharmaceutical focuses the immune response
CC towards selected determinants and could therefore be used in cases of
CC chronic viral diseases and cancer. Examples of diseases that can be
CC treated using the peptide containing pharmaceutical include autoimmune
CC diseases (rheumatoid arthritis, multiple sclerosis, and myasthenia
CC gravis), allograft rejection, allergies, lyme disease, hepatitis,
CC post-streptococcal endocarditis or glomerulonephritis and food
CC hypersensitivities. The peptide epitopes can be used to enhance immune
CC responses against other immunogens administered with the peptides.
CC Diseases which can be treated using immunogenic mixtures include prostate
CC cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma, cervical
CC carcinoma, lymphoma, and condyloma acuminatum. The peptides may also be
CC used to make monoclonal antibodies useful as potential diagnostic or
CC therapeutic agents. The peptides may also be useful as diagnostic
CC reagents, for example, to determine the susceptibility of an individual
CC to a treatment regimen. Also, the peptides may be used to predict which
CC individuals will be at substantial risk of developing chronic infection.
CC The selection of appropriate T and B cell epitopes should allow the
CC development of epitope based vaccines particularly towards conserved
CC epitopes of pathogens which are characterized by high sequence
CC variability such as HIV, HCV and Malaria.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 51; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VMAGVGSPYV 10
||||||||||
Db 5 vmagvgspyv 14

RESULT 7
Y46001
ID Y46001 standard; Peptide; 9 AA.
XX
AC Y46001;
XX
DT 01-DEC-1999 (first entry)
XX
DE Immunogenic peptide having a human leukocyte antigen binding motif #612.
XX
KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9945954-A1.
XX
PD 16-SEP-1999.
XX
PF 13-MAR-1998; 98WO-US05039.
XX
PR 13-MAR-1998; 98WO-US05039.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX
DR WPI; 1999-551214/46.
XX
PT New immunogenic peptides with HLA binding motif, useful in treatment
PT and diagnosis of cancers and viral diseases -
XX
PS Claim 1; Page 51; 150pp; English.
XX
CC Y45390 to Y48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes

```
CC   HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC   response against the antigen from which the peptide is derived.
CC   Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC   normally induced by an antigen in the form of a peptide fragment bound
CC   to a HLA molecule, rather than the intact foreign antigen itself, and
CC   are particularly important in tumour rejection and in fighting viral
CC   infections. The peptides are therefore useful therapeutically to treat
CC   or prevent viral infections and cancers in mammals (especially humans)
CC   e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC   They can be administered as vaccines to elicit an immune response in
CC   individuals susceptible or otherwise at risk of viral infection or
CC   cancer, or used to treat chronic or acute conditions. They are also
CC   useful diagnostically, and can be used to induce a cytotoxic T cell
CC   response, by contacting a cytotoxic T cell with the peptide e.g. to
CC   produce CTLs ex vivo for infusion back into a patient. The
CC   polynucleotides encoding the immunogenic peptides are also useful
CC   therapeutically and for immunisation as above.
XX
SQ   Sequence   9 AA;


  Query Match             92.2%;  Score 47;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VMAGVGSPY 9
          |||||||||
Db      1 vmagvgspy 9


RESULT   8
Y46413
ID   Y46413 standard; Peptide; 9 AA.
XX
AC   Y46413;
XX
DT   01-DEC-1999  (first entry)
XX
DE   Immunogenic peptide having a human leukocyte antigen binding motif #1024.
XX
KW   Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW   immune response; T cell activation; major histocompatibility complex;
KW   cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW   prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW   vaccine; immunisation.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9945954-A1.
XX
PD   16-SEP-1999.
XX
PF   13-MAR-1998;   98WO-US05039.
XX
PR   13-MAR-1998;   98WO-US05039.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Kubo RT,  Sidney J,  Celis E,  Grey HM,  Southwood S;
XX
DR   WPI; 1999-551214/46.
XX
PT   New immunogenic peptides with HLA binding motif, useful in treatment
PT   and diagnosis of cancers and viral diseases -
XX
PS   Claim 1; Page 71; 150pp; English.
XX
CC   Y45390 to Y48214 represent specifically claimed immunogenic peptides
CC   having a human major histocompatibility complex (MHC) Class I (also
CC   known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC   peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC   HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC   response against the antigen from which the peptide is derived.
CC   Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC   normally induced by an antigen in the form of a peptide fragment bound
CC   to a HLA molecule, rather than the intact foreign antigen itself, and
CC   are particularly important in tumour rejection and in fighting viral
CC   infections. The peptides are therefore useful therapeutically to treat
CC   or prevent viral infections and cancers in mammals (especially humans)
CC   e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC   They can be administered as vaccines to elicit an immune response in
CC   individuals susceptible or otherwise at risk of viral infection or
CC   cancer, or used to treat chronic or acute conditions. They are also
CC   useful diagnostically, and can be used to induce a cytotoxic T cell
CC   response, by contacting a cytotoxic T cell with the peptide e.g. to
CC   produce CTLs ex vivo for infusion back into a patient. The
CC   polynucleotides encoding the immunogenic peptides are also useful
CC   therapeutically and for immunisation as above.
XX
SQ   Sequence   9 AA;


  Query Match             92.2%;  Score 47;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VMAGVGSPY 9
          |||||||||
Db      1 vmagvgspy 9


RESULT   9
Y46478
ID   Y46478 standard; Peptide; 9 AA.
XX
AC   Y46478;
XX
DT   01-DEC-1999  (first entry)
XX
DE   Immunogenic peptide having a human leukocyte antigen binding motif #1089.
XX
KW   Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW   immune response; T cell activation; major histocompatibility complex;
KW   cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW   prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW   vaccine; immunisation.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9945954-A1.
XX
PD   16-SEP-1999.
XX
PF   13-MAR-1998;   98WO-US05039.
XX
PR   13-MAR-1998;   98WO-US05039.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Kubo RT,  Sidney J,  Celis E,  Grey HM,  Southwood S;
XX
DR   WPI; 1999-551214/46.
XX
PT   New immunogenic peptides with HLA binding motif, useful in treatment
PT   and diagnosis of cancers and viral diseases -
XX
PS   Claim 1; Page 74; 150pp; English.
XX
CC   Y45390 to Y48214 represent specifically claimed immunogenic peptides
CC   having a human major histocompatibility complex (MHC) Class I (also
CC   known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC   peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC   HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC   response against the antigen from which the peptide is derived.
```

CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in fighting viral
CC infections. The peptides are therefore useful therapeutically to treat
CC or prevent viral infections and cancers in mammals (especially humans)
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC They can be administered as vaccines to elicit an immune response in
CC individuals susceptible or otherwise at risk of viral infection or
CC cancer, or used to treat chronic or acute conditions. They are also
CC useful diagnostically, and can be used to induce a cytotoxic T cell
CC response, by contacting a cytotoxic T cell with the peptide e.g. to
CC produce CTLs ex vivo for infusion back into a patient. The
CC polynucleotides encoding the immunogenic peptides are also useful
CC therapeutically and for immunisation as above.
XX
SQ Sequence 9 AA;

Query Match 92.2%; Score 47; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VMAGVGSPY 9
|||||||||
Db 1 vmagvgspy 9

RESULT 10
Y37958
ID Y37958 standard; Peptide; 10 AA.
XX
AC Y37958;
XX
DT 29-SEP-1999 (first entry)
XX
DE Human cERB2 oncogene-derived HLA-binding peptide.
XX
KW Immunogen; HLA; human leukocyte antigen; binding motif; antiviral;
KW MHC; major histocompatability complex; viral infection; anticancer;
KW prostate cancer; lymphoma; hepatitis; AIDS; diagnostic; diagnosis.
XX
OS Homo sapiens.
XX
PN WO9403205-A1.
XX
PD 17-FEB-1994.
XX
PF 06-AUG-1993; 93WO-US07421.
XX
PR 05-MAR-1993; 93US-0027746.
PR 07-AUG-1992; 92US-0926666.
XX
PA (CYTE-) CYTEL CORP.
XX
PI Celis E, Grey HM, Kubo RT, Sette A;
XX
DR WPI; 1994-065403/08.
XX
PT Peptide which specifically binds selected MHC allele - used to
PT induce an immune response for treatment or prevention of viral
PT infection or cancer, or for diagnosis
XX
PS Disclosure; Page 103; 150pp; English.
XX
CC The sequence is a specific example of a group of new immunogenic
CC peptides having an HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1 binding
CC motif. For example, the peptides having an HLA-A3.2 binding motif
CC each have 9-10 residues and contain, from the N-terminus to the
CC C-terminus, (a) a first conserved residue selected from L, M, I,
CC V, S, A, T, F, C, G, D and E and (b) a second conserved residue of
CC K, R, Y, H or F, where the first and second conserved residues are
CC separated by 6-7 residues. The peptides are capable of binding
CC selected MHC molecules and inducing an immune response. They can be
CC used to treat and/or prevent viral infection and cancer, e.g. prostate
CC cancer, lymphoma, hepatitis or AIDS. They can also be used to produce
CC antibodies for use as diagnostic or therapeutic agents. The peptides
CC can also be used as diagnostic agents.
XX
SQ Sequence 10 AA;

Query Match 92.2%; Score 47; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0079;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VMAGVGSPY 9
|||||||||
Db 2 vmagvgspy 10

RESULT 11
Y45527
ID Y45527 standard; Peptide; 10 AA.
XX
AC Y45527;
XX
DT 01-DEC-1999 (first entry)
XX
DE Immunogenic peptide having a human leukocyte antigen binding motif #138.
XX
KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9945954-A1.
XX
PD 16-SEP-1999.
XX
PF 13-MAR-1998; 98WO-US05039.
XX
PR 13-MAR-1998; 98WO-US05039.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX
DR WPI; 1999-551214/46.
XX
PT New immunogenic peptides with HLA binding motif, useful in treatment
PT and diagnosis of cancers and viral diseases -
XX
PS Claim 1; Page 33; 150pp; English.
XX
CC Y45390 to Y48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC response against the antigen from which the peptide is derived.
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in fighting viral
CC infections. The peptides are therefore useful therapeutically to treat
CC or prevent viral infections and cancers in mammals (especially humans)
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC They can be administered as vaccines to elicit an immune response in
CC individuals susceptible or otherwise at risk of viral infection or
CC cancer, or used to treat chronic or acute conditions. They are also
CC useful diagnostically, and can be used to induce a cytotoxic T cell

```
CC   response, by contacting a cytotoxic T cell with the peptide e.g. to
CC   produce CTLs ex vivo for infusion back into a patient. The
CC   polynucleotides encoding the immunogenic peptides are also useful
CC   therapeutically and for immunisation as above.
XX
SQ   Sequence   10 AA;


  Query Match             92.2%;  Score 47;  DB 20;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.0079;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VMAGVGSPY 9
          |||||||||
Db      2 vmagvgspy 10


RESULT  12
Y98860
ID   Y98860 standard; Peptide; 15 AA.
XX
AC   Y98860;
XX
DT   07-AUG-2000  (first entry)
XX
DE   HLA class II binding antigen epitope peptide #49.
XX
KW   Human leucocyte antigen; HLA class II; antigen epitope; pharmaceutical;
KW   immune response; chronic viral disease; cancer; autoimmune disease;
KW   rheumatoid arthritis; multiple sclerosis; myasthenia gravis; AIDS;
KW   allograft rejection; allergy; lyme disease; hepatitis; prostate cancer;
KW   glomerulonephritis; food hypersensitivity; malaria.
XX
OS   Unidentified.
XX
PN   WO9961916-A1.
XX
PD   02-DEC-1999.
XX
PF   28-MAY-1999;   99WO-US12066.
XX
PR   29-MAY-1998;   98US-0087192.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Southwood S,  Sidney J;
XX
DR   WPI; 2000-097143/08.
XX
PT   New compositions containing immunogenic peptide epitopes for various
PT   HLA class II DR molecules useful for inducing helper T cell response
XX
PS   Claim 1; Page 40; 60pp; English.
XX
CC   The present invention relates to a new pharmaceutical composition
CC   comprising a unit dose form of a peptide, or analogue, comprising an
CC   epitope selected from those represented by peptides Y98812-Y99339 which
CC   are derived from various antigens for various human leucocyte antigen
CC   class DR molecules, representative of the world wide population. The
CC   peptide/analogue binds to an HLA class II molecule at an IC-50 of less
CC   than or equal to 1,000 nM. The pharmaceutical can be used to induce a
CC   helper T cell response. The pharmaceutical focuses the immune response
CC   towards selected determinants and could therefore be used in cases of
CC   chronic viral diseases and cancer. Examples of diseases that can be
CC   treated using the peptide containing pharmaceutical include autoimmune
CC   diseases (rheumatoid arthritis, multiple sclerosis, and myasthenia
CC   gravis), allograft rejection, allergies, lyme disease, hepatitis,
CC   post-streptococcal endocarditis or glomerulonephritis and food
CC   hypersensitivities. The peptide epitopes can be used to enhance immune
CC   responses against other immunogens administered with the peptides.
CC   Diseases which can be treated using immunogenic mixtures include prostate
CC   cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma, cervical
CC   carcinoma, lymphoma, and condyloma acuminatum. The peptides may also be
CC   used to make monoclonal antibodies useful as potential diagnostic or
CC   therapeutic agents. The peptides may also be useful as diagnostic
CC   reagents, for example, to determine the susceptibility of an individual
CC   to a treatment regimen. Also, the peptides may be used to predict which
CC   individuals will be at substantial risk of developing chronic infection.
CC   The selection of appropriate T and B cell epitopes should allow the
CC   development of epitope based vaccines particularly towards conserved
CC   epitopes of pathogens which are characterized by high sequence
CC   variability such as HIV, HCV and Malaria.
XX
SQ   Sequence   15 AA;


  Query Match             64.7%;  Score 33;  DB 21;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 4.7;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VMAGVGS 7
          |||||||
Db      9 vmagvgs 15


RESULT  13
W30834
ID   W30834 standard; Peptide; 9 AA.
XX
AC   W30834;
XX
DT   20-MAR-1998  (first entry)
XX
DE   TRP-2 derived potential cancer antigen 7, based on positions 197-205.
XX
KW   Tyrosinase related protein 2 gene; TRP-1; gp75; tumour antigen;
KW   tumour infiltrating lymphocyte; TIL; TIL586; cancer peptide; TRP-2;
KW   alternative reading frame; cancer detection; pre-cancer detection;
KW   melanoma.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /label= L2A
FT                   /note= "wild type Leu198 substituted with Ala"
XX
PN   WO9729195-A2.
XX
PD   14-AUG-1997.
XX
PF   06-FEB-1997;   97WO-US02186.
XX
PR   04-OCT-1996;   96US-0725736.
PR   09-FEB-1996;   96US-0599602.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Rosenberg SA,  Wang R;
XX
DR   WPI; 1997-415349/38.
XX
PT   Cancer antigen peptide(s) derived from the tyrosinase-related
PT   protein 1 or 2  - useful for detecting, preventing or treating a
PT   cancer in a mammal, especially melanoma
XX
PS   Claim 17; Page 56; 111pp; English.
XX
CC   Peptides W30829-38 and W37011-21 are modified versions of a peptide
CC   derived from positions 197-205 of the tyrosinase related protein 2
CC   (TRP-2). This region contains the peptide epitope of TRP-2 that is
CC   able to stimulate cytokine release by CTL cells. Apart from
CC   W30829-30 (these contain extra residues at the N-terminal (W30829) and
CC   the C-terminal (W30830)), the peptides were modified to contain
CC   substitutions at the anchor residues. Of all these peptides, only
```

CC W30829-38 were able to stimulate cytokine release. Other antigenic
CC peptides have also been identified from TRP-1. The nucleic acids encoding
CC the cancer peptides or TRP-2 can be used to detect a cancer or pre-cancer
CC in a mammal, especially by detecting the presence of the alterative ORF 3
CC of the TRP-1 gene or the sequence encoding the novel tumour antigen
CC TRP-2. Vectors and recombinant viruses containing antigen peptide
CC encoding nucleic acids, antibodies raised against the peptides, or the
CC peptides themselves can be used to prevent or treat a cancer in a mammal,
CC especially a melanoma.
XX
SQ Sequence 9 AA;

Query Match 56.9%; Score 29; DB 18; Length 9;
Best Local Similarity 62.5%; Pred. No. 2.1e+05;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MAGVGSPY 9
:|| | ||
Db 1 lagpgrpy 8

RESULT 14
Y98957
ID Y98957 standard; Peptide; 15 AA.
XX
AC Y98957;
XX
DT 07-AUG-2000 (first entry)
XX
DE HLA class II binding antigen epitope peptide #146.
XX
KW Human leucocyte antigen; HLA class II; antigen epitope; pharmaceutical;
KW immune response; chronic viral disease; cancer; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; myasthenia gravis; AIDS;
KW allograft rejection; allergy; lyme disease; hepatitis; prostate cancer;
KW glomerulonephritis; food hypersensitivity; malaria.
XX
OS Unidentified.
XX
PN WO9961916-A1.
XX
PD 02-DEC-1999.
XX
PF 28-MAY-1999; 99WO-US12066.
XX
PR 29-MAY-1998; 98US-0087192.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Southwood S, Sidney J;
XX
DR WPI; 2000-097143/08.
XX
PT New compositions containing immunogenic peptide epitopes for various
PT HLA class II DR molecules useful for inducing helper T cell response
XX
PS Claim 1; Page 42; 60pp; English.
XX
CC The present invention relates to a new pharmaceutical composition
CC comprising a unit dose form of a peptide, or analogue, comprising an
CC epitope selected from those represented by peptides Y98812-Y99339 which
CC are derived from various antigens for various human leucocyte antigen
CC class DR molecules, representative of the world wide population. The
CC peptide/analogue binds to an HLA class II molecule at an IC-50 of less
CC than or equal to 1,000 nM. The pharmaceutical can be used to induce a
CC helper T cell response. The pharmaceutical focuses the immune response
CC towards selected determinants and could therefore be used in cases of
CC chronic viral diseases and cancer. Examples of diseases that can be
CC treated using the peptide containing pharmaceutical include autoimmune
CC diseases (rheumatoid arthritis, multiple sclerosis, and myasthenia
CC gravis), allograft rejection, allergies, lyme disease, hepatitis,
CC post-streptococcal endocarditis or glomerulonephritis and food
CC hypersensitivities. The peptide epitopes can be used to enhance immune
CC responses against other immunogens administered with the peptides.
CC Diseases which can be treated using immunogenic mixtures include prostate
CC cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma, cervical
CC carcinoma, lymphoma, and condyloma acuminatum. The peptides may also be
CC used to make monoclonal antibodies useful as potential diagnostic or
CC therapeutic agents. The peptides may also be useful as diagnostic
CC reagents, for example, to determine the susceptibility of an individual
CC to a treatment regimen. Also, the peptides may be used to predict which
CC individuals will be at substantial risk of developing chronic infection.
CC The selection of appropriate T and B cell epitopes should allow the
CC development of epitope based vaccines particularly towards conserved
CC epitopes of pathogens which are characterized by high sequence
CC variability such as HIV, HCV and Malaria.
XX
SQ Sequence 15 AA;

Query Match 56.9%; Score 29; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VMAGVG 6
||||||
Db 10 vmagvg 15

RESULT 15
W30833
ID W30833 standard; Peptide; 9 AA.
XX
AC W30833;
XX
DT 20-MAR-1998 (first entry)
XX
DE TRP-2 derived potential cancer antigen 6, based on positions 197-205.
XX
KW Tyrosinase related protein 2 gene; TRP-1; gp75; tumour antigen;
KW tumour infiltrating lymphocyte; TIL; TIL586; cancer peptide; TRP-2;
KW alternative reading frame; cancer detection; pre-cancer detection;
KW melanoma.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /label= L2S
FT /note= "wild type Leu198 substituted with Ser"
XX
PN WO9729195-A2.
XX
PD 14-AUG-1997.
XX
PF 06-FEB-1997; 97WO-US02186.
XX
PR 04-OCT-1996; 96US-0725736.
PR 09-FEB-1996; 96US-0599602.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Rosenberg SA, Wang R;
XX
DR WPI; 1997-415349/38.
XX
PT Cancer antigen peptide(s) derived from the tyrosinase-related
PT protein 1 or 2 - useful for detecting, preventing or treating a
PT cancer in a mammal, especially melanoma
XX
PS Claim 17; Page 56; 111pp; English.
XX
CC Peptides W30829-38 and W37011-21 are modified versions of a peptide
CC derived from positions 197-205 of the tyrosinase related protein 2

```
CC   (TRP-2). This region contains the peptide epitope of TRP-2 that is
CC   able to stimulate cytokine release by CTL cells. Apart from
CC   W30829-30 (these contain extra residues at the N-terminal (W30829) and
CC   the C-terminal (W30830)), the peptides were modified to contain
CC   substitutions at the anchor residues. Of all these peptides, only
CC   W30829-38 were able to stimulate cytokine release. Other antigenic
CC   peptides have also been identified from TRP-1. The nucleic acids encoding
CC   the cancer peptides or TRP-2 can be used to detect a cancer or pre-cancer
CC   in a mammal, especially by detecting the presence of the alterative ORF 3
CC   of the TRP-1 gene or the sequence encoding the novel tumour antigen
CC   TRP-2. Vectors and recombinant viruses containing antigen peptide
CC   encoding nucleic acids, antibodies raised against the peptides, or the
CC   peptides themselves can be used to prevent or treat a cancer in a mammal,
CC   especially a melanoma.
XX
SQ   Sequence   9 AA;


  Query Match             51.0%;  Score 26;  DB 18;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.1e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 MAGVGSPY 9
          ::| | ||
Db      1 lsgpgrpy 8


RESULT  16
W37019
ID   W37019 standard; Peptide; 9 AA.
XX
AC   W37019;
XX
DT   20-MAR-1998  (first entry)
XX
DE   TRP-2 derived potential cancer antigen 20, based on positions 197-205.
XX
KW   Tyrosinase related protein 2 gene; TRP-1; gp75; tumour antigen;
KW   tumour infiltrating lymphocyte; TIL; TIL586; cancer peptide; TRP-2;
KW   alternative reading frame; cancer detection; pre-cancer detection;
KW   melanoma.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 6
FT                   /label= R6A
FT                   /note= "wild type Arg202 substituted with Ala"
XX
PN   WO9729195-A2.
XX
PD   14-AUG-1997.
XX
PF   06-FEB-1997;   97WO-US02186.
XX
PR   04-OCT-1996;   96US-0725736.
PR   09-FEB-1996;   96US-0599602.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Rosenberg SA,  Wang R;
XX
DR   WPI; 1997-415349/38.
XX
PT   Cancer antigen peptide(s) derived from the tyrosinase-related
PT   protein 1 or 2  - useful for detecting, preventing or treating a
PT   cancer in a mammal, especially melanoma
XX
PS   Example 14; Page 56; 111pp; English.
XX
CC   Peptides W30829-38 and W37011-21 are modified versions of a peptide
CC   derived from positions 197-205 of the tyrosinase related protein 2
```


```
CC   (TRP-2). This region contains the peptide epitope of TRP-2 that is
CC   able to stimulate cytokine release by CTL cells. Apart from
CC   W30829-30 (these contain extra residues at the N-terminal (W30829) and
CC   the C-terminal (W30830)), the peptides were modified to contain
CC   substitutions at the anchor residues. Of all these peptides, only
CC   W30829-38 were able to stimulate cytokine release. Other antigenic
CC   peptides have also been identified from TRP-1. The nucleic acids encoding
CC   the cancer peptides or TRP-2 can be used to detect a cancer or pre-cancer
CC   in a mammal, especially by detecting the presence of the alterative ORF 3
CC   of the TRP-1 gene or the sequence encoding the novel tumour antigen
CC   TRP-2. Vectors and recombinant viruses containing antigen peptide
CC   encoding nucleic acids, antibodies raised against the peptides, or the
CC   peptides themselves can be used to prevent or treat a cancer in a mammal,
CC   especially a melanoma.
XX
SQ   Sequence   9 AA;


  Query Match             51.0%;  Score 26;  DB 18;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.1e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;
```


```
Qy      2 MAGVGSPY 9
          : | |:||
Db      1 llgpgapy 8


RESULT  17
W38105
ID   W38105 standard; peptide; 16 AA.
XX
AC   W38105;
XX
DT   23-APR-1998  (first entry)
XX
DE   Peptide recognition unit WBP-2B used to identify WW domains.
XX
KW   Peptide recognition unit; YAP WW domain binding protein; WBP-1; WBP-2;
KW   WW domain; cell signalling; growth regulation; cytoskeleton organisation;
KW   targeted drug screening; modulator; WW domain interaction.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "residue is biotinylated"
XX
PN   WO9737223-A1.
XX
PD   09-OCT-1997.
XX
PF   03-APR-1997;   97WO-US05547.
XX
PR   03-APR-1996;   96US-0630916.
XX
PA   (CYTO-) CYTOGEN CORP.
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Fowlkes DM,  Kay BK,  Pirozzi G;
XX
DR   WPI; 1997-503234/46.
XX
PT   Identifying cell signalling and growth regulatory polypeptides by
PT   reaction with multivalent recognition complex - polypeptides are
PT   useful in targetted drug selection
XX
PS   Example 1; Page 70; 220pp; English.
XX
CC   Peptides W38103-05 are peptide recognition units that are based on the
CC   sequences of the YAP WW domain binding proteins WBP-1 and WBP-2. They
CC   were used to screen human bone marrow and brain cDNA libraries. 13
CC   cDNA clones were identified and isolated. These clones represented 3
CC   novel human genes WWP1 (W36794), WWP2 (W36795) and WWP3 (W36796). The
```

```
CC   WW domain is a small functional domain found in a large number of
CC   proteins from a variety of species including humans, nematodes and yeast.
CC   Its name is derived from the observation that two tryptophan residues,
CC   one in the amino terminal portion of the WW domain and one in the
CC   carboxyl terminal portion, are conserved. Most proteins containing WW
CC   domains have a function involving cell signalling and growth regulation
CC   or the organisation of the cytoskeleton. Polypeptides containing a WW
CC   domain are identified by treating a multivalent recognition unit complex
CC   that has selective binding affinity for a WW domain, with many
CC   polypeptides and identifying those with selective affinity for the
CC   complex. Proteins containing WW domains are used for targeted drug
CC   screening, i.e. to identify potential modulators of specific WW domain
CC   interactions. The valency of the recognition unit is important in
CC   determining specificity of interaction with WW domains. In multivalent
CC   form specificity is relaxed, but not lost, so proteins containing WW
CC   domains similar, but not identical, to the sequence of the peptides'
CC   target WW can be detected, including new polypeptides.
XX
SQ   Sequence   16 AA;


  Query Match             51.0%;  Score 26;  DB 18;  Length 16;
  Best Local Similarity   57.1%;  Pred. No. 1e+02;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 AGVGSPY 9
          :| |:||
Db      3 sgpgtpy 9


RESULT  18
Y07541
ID   Y07541 standard; peptide; 6 AA.
XX
AC   Y07541;
XX
DT   25-APR-2000  (first entry)
XX
DE   Hexapeptide having antiarrhythmic activity.
XX
KW   Antiarrhythmic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   4
FT                   /label= 4Hyp
FT                   /note= "4-hydroxyproline residue"
FT   Modified-site   6
FT                   /note= "3-iodo-tyrosinamide (Claim 2); also
FT                    3-fluoro-, 3 -chloro- or 3-bromo-tyrosinamide in
FT                    Examples (Page 4)"
XX
PN   WO9621674-A1.
XX
PD   18-JUL-1996.
XX
PF   04-JAN-1996;   96WO-EP00009.
XX
PR   14-JAN-1995;   95DE-1000990.
XX
PA   (BADI ) BASF AG.
XX
PI   Dhein S,  Tudyka T;
XX
DR   WPI; 1996-342238/34.
XX
PT   New hexapeptide derivs. having antiarrhythmic activity - useful for
PT   preventing ischaemia- or age-related rhythm disorders without
PT   significant pro-arrhythmic activity
XX
PS   Claim 2; Page 7; 13pp; German.
XX
CC   The patent discloses new hexapeptide compounds having antiarrhythmic
CC   activity which are useful for preventing ischaemia- or age-related
CC   rhythm disorders without significant pro-arrhythmic activity. They
CC   suppress local differences in the duration of action potential and
CC   irregularities in the spread of the exciting stimulus. The peptides
CC   have the generic formula H2N-X-Ala-Gly-Hyp-Y-Z-NH2 (see Y07540) in
CC   which X is Ala, Arg, Gly or Val, Y is Pro or His, and Z is Tyr or Phe
CC   which is optionally ring-substituted by I, F, Cl or Br.  The present
CC   sequence represents a preferred example of the new peptides.
XX
SQ   Sequence   6 AA;


  Query Match             49.0%;  Score 25;  DB 17;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GVGSPY 9
          | | ||
Db      1 gagppy 6


RESULT  19
W11180
ID   W11180 standard; peptide; 6 AA.
XX
AC   W11180;
XX
DT   09-NOV-1998  (first entry)
XX
DE   Cyclic antiarrhythmic peptide sequence.
XX
KW   Antiarrhythmic; cyclic; AAP10.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1..6
FT                   /note= "Gly(1) and Tyr(6) are condensed via the
FT                    linkage (Tyr6)-CONH-C(CF3)(OH)-(Gly1) or
FT                    (Tyr6)-CONH-CO-(Gly1), thus forming a cyclic peptide"
FT   Modified-site   4
FT                   /label= 4Hyp
XX
PN   DE19707854-A1.
XX
PD   03-SEP-1998.
XX
PF   27-FEB-1997;   97DE-1007854.
XX
PR   27-FEB-1997;   97DE-1007854.
XX
PA   (DHEI/) DHEIN S.
PA   (GROV/) GROVER R.
XX
PI   Dhein S,  Grover R;
XX
DR   WPI; 1998-468293/41.
XX
PT   New cyclic peptides used as antiarrhythmic agents - are more stable
PT   than linear antiarrhythmic peptides
XX
PS   Claims 1, 3; Page 4; 4pp; German.
XX
CC   The present sequence represents the new cyclic peptides
CC   cyclo[CF3C(OH)-Gly-Ala-Gly-4-Hyp-Pro-Tyr-CONH] and
CC   cyclo[CO-Gly-Ala-Gly-4-Hyp-Pro-Tyr-CONH], which are useful as
CC   antiarrhythmic agents. They have better stability in solution than
CC   the known corresponding linear antiarrhythmic peptide AAP10, i.e.
CC   NH2-Gly-Ala-Gly-4-Hyp-Pro-Tyr-CONH2.
XX
SQ   Sequence   6 AA;
```

```
  Query Match             49.0%;  Score 25;  DB 19;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GVGSPY 9
          | | ||
Db      1 gagppy 6


RESULT  20
R95514
ID   R95514 standard; peptide; 8 AA.
XX
AC   R95514;
XX
DT   06-NOV-1996  (first entry)
XX
DE   Human prostate carcinoma cell antigen binding peptide ("abtide").
XX
KW   Abtide; prostate specific mucin antigen; human prostate cancer; LNCaP;
KW   diagnostic; detection; imaging; tumour; phage; peptide library;
KW   polymorphic; epithelial.
XX
OS   Synthetic.
XX
PN   WO9609411-A1.
XX
PD   28-MAR-1996.
XX
PF   20-SEP-1995;   95WO-US11934.
XX
PR   07-JUN-1995;   95US-0488161.
PR   21-SEP-1994;   94US-0310192.
XX
PA   (CYTO-) CYTOGEN CORP.
XX
PI   Alvarez VL;
XX
DR   WPI; 1996-188471/19.
XX
PT   New isolated peptide(s) with specific binding activities - obtd. by
PT   screening random peptide libraries, for use in diagnostic and
PT   therapeutic compsns.
XX
PS   Claim 48; Page 93; 106pp; English.
XX
CC   R95511-R95520 are antigen binding peptides ("abtides"), which bind to
CC   a human prostate carcinoma cell antigen. The abtides are identified
CC   from random peptide libraries using specific ligand binding. Abtides
CC   mimic the binding specificity of large molecules such as antibodies
CC   and receptors but have a much smaller size allowing their production
CC   at a lower cost and reducing the extent of their immunogenicity aiding
CC   in vivo delivery. The abtides are useful for the diagnosis, detection,
CC   imaging and treatment of disease, e.g. tumours, prostate cancer.
XX
SQ   Sequence   8 AA;


  Query Match             49.0%;  Score 25;  DB 17;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 GVGSPY 9
          || :||
Db      3 gvstpy 8


RESULT  21
Y59138
ID   Y59138 standard; peptide; 8 AA.
XX
AC   Y59138;
XX
DT   08-MAR-2000  (first entry)
XX
DE   Human PAK65 kinase domain conserved fragment.
XX
KW   PAK4; serine/threonine kinase; GTPase; intracellular signal cascade;
KW   Rac; Cdc42H; morphogenesis; mitogenesis; JNK; p38 MAP kinase; human;
KW   actin polymerization; filopodia; cancer; arthritis; PAK65; STE20.
XX
OS   Saccharomyces cerevisiae.
XX
PN   WO9963073-A1.
XX
PD   09-DEC-1999.
XX
PF   21-MAY-1999;   99WO-US11341.
XX
PR   21-MAY-1998;   98US-0082737.
XX
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Minden A;
XX
DR   WPI; 2000-072881/06.
XX
PT   Novel mammalian nucleic acid useful for treating cancer and arthritis
PT   -
XX
PS   Disclosure; Page 29; 95pp; English.
XX
CC   The invention relates to an isolated mammalian nucleic acid that encodes
CC   PAK4, a novel serine/threonine kinase or its mutant homolog. PAK4 is an
CC   effector for the GTPases Rac and Cdc42Hs which are involved in
CC   intracellular signal cascades, morphogenesis and mitogenesis, and
CC   activate the JNK and p38 MAP kinase pathways. Inhibiting interaction of
CC   PAK4 with these enzymes will thus result in inhibition of actin
CC   polymerization and formation of filopodia. The PAK4 nucleic acid used for
CC   recombinant production of the protein, and as a source of probes for
CC   identifying homologous sequences and of (anti)sense oligonucleotides for
CC   inhibiting PAK4 expression. The protein, or its fragments, are used to
CC   raise specific antibodies and these are useful as ligands for therapeutic
CC   inhibition of interaction between PAK4 and its native binding partners.
CC   Inhibition of PAK4 activity or expression is used for treatment of cancer
CC   and arthritis. The present sequence represents a conserved kinase domain
CC   fragment of human PAK65 and yeast STE20. This is used to design
CC   degenerate primers for isolation of human PAK4.
XX
SQ   Sequence   8 AA;


  Query Match             49.0%;  Score 25;  DB 21;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 VGSPY 9
          ||:||
Db      1 vgtpy 5


RESULT  22
W30838
ID   W30838 standard; Peptide; 9 AA.
XX
AC   W30838;
XX
DT   20-MAR-1998  (first entry)
XX
DE   TRP-2 derived potential cancer antigen 11, based on positions 197-205.
XX
KW   Tyrosinase related protein 2 gene; TRP-1; gp75; tumour antigen;
KW   tumour infiltrating lymphocyte; TIL; TIL586; cancer peptide; TRP-2;
KW   alternative reading frame; cancer detection; pre-cancer detection;
KW   melanoma.
```

```
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 6
FT                   /label= R6K
FT                   /note= "wild type Arg202 substituted with Lys"
XX
PN   WO9729195-A2.
XX
PD   14-AUG-1997.
XX
PF   06-FEB-1997;   97WO-US02186.
XX
PR   04-OCT-1996;   96US-0725736.
PR   09-FEB-1996;   96US-0599602.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Rosenberg SA,  Wang R;
XX
DR   WPI; 1997-415349/38.
XX
PT   Cancer antigen peptide(s) derived from the tyrosinase-related
PT   protein 1 or 2  - useful for detecting, preventing or treating a
PT   cancer in a mammal, especially melanoma
XX
PS   Claim 17; Page 56; 111pp; English.
XX
CC   Peptides W30829-38 and W37011-21 are modified versions of a peptide
CC   derived from positions 197-205 of the tyrosinase related protein 2
CC   (TRP-2). This region contains the peptide epitope of TRP-2 that is
CC   able to stimulate cytokine release by CTL cells. Apart from
CC   W30829-30 (these contain extra residues at the N-terminal (W30829) and
CC   the C-terminal (W30830)), the peptides were modified to contain
CC   substitutions at the anchor residues. Of all these peptides, only
CC   W30829-38 were able to stimulate cytokine release. Other antigenic
CC   peptides have also been identified from TRP-1. The nucleic acids encoding
CC   the cancer peptides or TRP-2 can be used to detect a cancer or pre-cancer
CC   in a mammal, especially by detecting the presence of the alterative ORF 3
CC   of the TRP-1 gene or the sequence encoding the novel tumour antigen
CC   TRP-2. Vectors and recombinant viruses containing antigen peptide
CC   encoding nucleic acids, antibodies raised against the peptides, or the
CC   peptides themselves can be used to prevent or treat a cancer in a mammal,
CC   especially a melanoma.
XX
SQ   Sequence   9 AA;


  Query Match            49.0%;  Score 25;  DB 18;  Length 9;
  Best Local Similarity  50.0%;  Pred. No. 2.1e+05;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MAGVGSPY 9
           : | | ||
Db       1 llgpgkpy 8


RESULT  23
W30832
ID   W30832 standard; Peptide; 9 AA.
XX
AC   W30832;
XX
DT   20-MAR-1998  (first entry)
XX
DE   TRP-2 derived potential cancer antigen 5, based on positions 197-205.
XX
KW   Tyrosinase related protein 2 gene; TRP-1; gp75; tumour antigen;
KW   tumour infiltrating lymphocyte; TIL; TIL586; cancer peptide; TRP-2;
KW   alternative reading frame; cancer detection; pre-cancer detection;
KW   melanoma.
```

```
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /label= L2V
FT                   /note= "wild type Leu198 substituted with Val"
XX
PN   WO9729195-A2.
XX
PD   14-AUG-1997.
XX
PF   06-FEB-1997;   97WO-US02186.
XX
PR   04-OCT-1996;   96US-0725736.
PR   09-FEB-1996;   96US-0599602.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Rosenberg SA,  Wang R;
XX
DR   WPI; 1997-415349/38.
XX
PT   Cancer antigen peptide(s) derived from the tyrosinase-related
PT   protein 1 or 2  - useful for detecting, preventing or treating a
PT   cancer in a mammal, especially melanoma
XX
PS   Claim 17; Page 56; 111pp; English.
XX
CC   Peptides W30829-38 and W37011-21 are modified versions of a peptide
CC   derived from positions 197-205 of the tyrosinase related protein 2
CC   (TRP-2). This region contains the peptide epitope of TRP-2 that is
CC   able to stimulate cytokine release by CTL cells. Apart from
CC   W30829-30 (these contain extra residues at the N-terminal (W30829) and
CC   the C-terminal (W30830)), the peptides were modified to contain
CC   substitutions at the anchor residues. Of all these peptides, only
CC   W30829-38 were able to stimulate cytokine release. Other antigenic
CC   peptides have also been identified from TRP-1. The nucleic acids encoding
CC   the cancer peptides or TRP-2 can be used to detect a cancer or pre-cancer
CC   in a mammal, especially by detecting the presence of the alterative ORF 3
CC   of the TRP-1 gene or the sequence encoding the novel tumour antigen
CC   TRP-2. Vectors and recombinant viruses containing antigen peptide
CC   encoding nucleic acids, antibodies raised against the peptides, or the
CC   peptides themselves can be used to prevent or treat a cancer in a mammal,
CC   especially a melanoma.
XX
SQ   Sequence   9 AA;


  Query Match            49.0%;  Score 25;  DB 18;  Length 9;
  Best Local Similarity  50.0%;  Pred. No. 2.1e+05;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MAGVGSPY 9
           : | | ||
Db       1 lvgpgrpy 8


RESULT  24
R57426
ID   R57426 standard; Protein; 12 AA.
XX
AC   R57426;
XX
DT   14-MAR-1995  (first entry)
XX
DE   Rabphilin-3A fragment, peak 5.
XX
KW   Low molecular weight; G protein; target protein; rab3A p25;
KW   Rabphilin-3A; brain; nerve transmitter.
XX
OS   Homo sapiens.
```

```
XX
PN   JP06184199-A.
XX
PD   05-JUL-1994.
XX
PF   24-DEC-1992;   92JP-0344055.
XX
PR   24-DEC-1992;   92JP-0344055.
XX
PA   (EISA ) EISAI CO LTD.
XX
DR   WPI; 1994-252836/31.
XX
PT   Target protein of a low molecular G protein rabphilin-3A (RAB3A)
PT   - found in the brain and involved in release of nerve transmitter
PT   substance
XX
PS   Example 3; Page 5; 9pp; Japanese.
XX
CC   The sequences given in R57422-32 represents fragments of low molecular
CC   weight G protein target protein, designated rab3A p25.  Rabphilin-3A
CC   (Rab3A) is distributed specifically in brain tissue and participates
CC   in the release of nerve transmitter substance and is useful in the study
CC   of its secretion.
XX
SQ   Sequence   12 AA;


  Query Match             49.0%;  Score 25;  DB 15;  Length 12;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GVGSPYV 10
           |:  |||
Db       6 gladpyv 12


RESULT  25
W62130
ID   W62130 standard; peptide; 12 AA.
XX
AC   W62130;
XX
DT   16-SEP-1998  (first entry)
XX
DE   Haemophilus influenzae tyrosine tRNA synthetase binding peptide 3.
XX
KW   Identification; ligand; biological activity; target-binding;
KW   drug screening; library; inhibitory ligand.
XX
OS   Synthetic.
OS   Haemophilus influenzae.
XX
PN   WO9819162-A1.
XX
PD   07-MAY-1998.
XX
PF   31-OCT-1997;   97WO-US19638.
XX
PR   31-OCT-1996;   96US-0740671.
XX
PA   (NOVA-) NOVALON PHARM CORP.
XX
PI   Fowlkes DM,  Frelinger JA,  Hyde-Deruyscher RP,  Kay BK;
XX
DR   WPI; 1998-272389/24.
XX
PT   Identifying ligands which mediate biological activity of a protein -
PT   by identifying target-binding ligands and screening a library for
PT   ligands which inhibit target-binding ligand mediated activity
XX
PS   Example 5; Page 100; 143pp; English.
XX
CC   A method has been developed for identifying a ligand which mediates the
CC   biological activity of a target protein (T) by inhibiting the binding
CC   of (T) to a binding partner. The method comprises: (a) screening a first
CC   combinatorial library comprising first member ligands for binding to the
CC   target-binding ligands (TBLs), to identifying one or more TBLs; (b)
CC   screening a second library comprising second member ligands for the
CC   ability to inhibit the binding of one or more of the TBLs to the target
CC   protein, and so obtaining one or more inhibitory ligands; and (c)
CC   determining which of the inhibitory ligands can mediate a biological
CC   activity of the target protein. The present sequence represents a
CC   potential binding peptide for Haemophilus influenzae tyrosine tRNA
CC   synthetase from an example of the present invention. The method can be
CC   used for identifying drugs which can mediate the biological activity of
CC   a target protein. It can be used to identify the biological activity of
CC   a target protein whose biological function is not known and perhaps
CC   cannot be determined directly. The method can also be used to identify
CC   new inhibitory ligands of specific target proteins. The method provides
CC   high throughput screens which are essentially identical for similar and
CC   dissimilar targets, bypassing the need to develop distinct assays for
CC   biochemically diverse targets.
XX
SQ   Sequence   12 AA;


  Query Match             49.0%;  Score 25;  DB 19;  Length 12;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 GSPYV 10
           ||||:
Db       6 gspyl 10


RESULT  26
W15779
ID   W15779 standard; Peptide; 13 AA.
XX
AC   W15779;
XX
DT   31-OCT-1997  (first entry)
XX
DE   Protein kinase C-beta peptide beta C2-2 (186-198).
XX
KW   Signal transduction; cell signalling; modulator; immunomodulator;
KW   protein kinase C; receptor for activated kinase C; RACK; PKC-beta;
KW   cognate; graft rejection; autoimmune disease; allergy; asthma;
KW   therapy.
XX
OS   Homo sapiens.
XX
PN   WO9714038-A1.
XX
PD   17-APR-1997.
XX
PF   10-OCT-1996;   96WO-US16195.
XX
PR   18-JUN-1996;   96US-0665647.
PR   10-OCT-1995;   95US-0541964.
PR   31-JAN-1996;   96US-0594447.
XX
PA   (TERR-) TERRAPIN TECHNOLOGIES INC.
XX
PI   Kauvar LM,  Mochly-Rosen D,  Napolitano EW,  Ron D;
PI   Vasquez NJ,  Voronova A;
XX
DR   WPI; 1997-236030/21.
XX
PT   Identifying a modulator of intracellular signal transduction - by
PT   determining the interaction of a signal generating peptide with the
PT   test substance, allows modulation of the immune system
XX
PS   Claim 9; Page 26; 74pp; English.
XX
```

```
CC   This sequence is a peptide, designated beta C2-2, that corresponds
CC   to amino acid residues 186-198 in the C2 region of protein kinase
CC   C-beta (PKC-beta).  It is capable of interrupting the interaction
CC   of PKC-beta with its cognate receptor for activated kinase C
CC   (RACK1).  Beta C2-2 can be used as a signal generating peptide in a
CC   claimed method for identifying modulators of intracellular signal
CC   transduction.  This method assesses the abillty of candidate
CC   modulators to affect the interaction between a signal-generating
CC   protein, such as a PKC isozyme peptide (see also W15778, W15781,
CC   W15784-85, W17452-78), and a cognate binding protein involved in
CC   modulating the signal transduction function.  Identified substances
CC   are useful as immunomodulators (claimed).  They act to reduce T-cell
CC   activity, reduce the rate of graft rejection, reduce the severity of
CC   an autoimmune disorder, ameliorate allergy and/or asthma, or
CC   diminish a cytokine response (claimed).
XX
SQ   Sequence   13 AA;


  Query Match             49.0%;  Score 25;  DB 18;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GVGSPYV 10
           |:  |||
Db       5 glsdpyv 11


RESULT  27
Y66838
ID   Y66838 standard; peptide; 13 AA.
XX
AC   Y66838;
XX
DT   11-APR-2000  (first entry)
XX
DE   T cell antigen receptor Vbeta 4 chain peptide.
XX
KW   Rheumatoid arthritis; arthrosis deformans; T-cell antigen receptor;
KW   Vbeta chain; autoantigen; immunological tolerance.
XX
OS   Homo sapiens.
XX
PN   WO9963084-A1.
XX
PD   09-DEC-1999.
XX
PF   28-MAY-1999;   99WO-JP02814.
XX
PR   29-MAY-1998;   98JP-0149855.
PR   14-OCT-1998;   98JP-0328761.
XX
PA   (TORI-) TORII PHARM CO LTD.
XX
PI   Nishioka K,  Yoshino S;
XX
DR   WPI; 2000-086978/07.
DR   N-PSDB; Z96568.
XX
PT   T-cell antigen receptor V-beta chain CDR3 region sequences accumulated
PT   in synovial membranes of rheumatoid arthritis patients  -
XX
PS   Example 3; Page 59; 136pp; Japanese.
XX
CC   The invention relates to peptide sequences present in the synovial fluid
CC   and membranes of rheumatoid arthritis patients, arising from the CDR
CC   region of oligoclonal pathogenic T-cell antigen receptor Vbeta chains.
CC   Compositions which contain autoantigenic peptides binding specifically
CC   to T-cells expressing receptors containing the peptide sequences, which
CC   include antigen-specific immunological tolerance to rheumatoid arthritis
CC   can be used for the treatment and prevention of rheumatoid arthritis.
CC   The invention can be used for the diagnosis, treatment and prevention
CC   of rheumatoid arthritis. Sequences Y66771-958 represent peptides from
CC   the various Vbeta chains of T cell antigen receptor.
XX
SQ   Sequence   13 AA;


  Query Match             49.0%;  Score 25;  DB 21;  Length 13;
  Best Local Similarity   55.6%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 VMAGVGSPY 9
           |  |||: |
Db       4 vptgvgtqy 12


RESULT  28
W85388
ID   W85388 standard; peptide; 15 AA.
XX
AC   W85388;
XX
DT   16-FEB-1999  (first entry)
XX
DE   Helper T-cell class II peptide derived from SSP2 protein.
XX
KW   Helper T-cell peptide; human leucocyte antigen; HLA; DR4w4; DR1;
KW   DR7; cytotoxic T lymphocyte; CTL; hepatitis; autoimmune disease;
KW   acquired immune deficiency syndrome; malaria; cancer;
KW   allograft rejection; allergy; Lyme disease; hepatitis;
KW   post-streptococcal endocarditis; glomerulonephritis;
KW   food hypersensitivity.
XX
OS   Synthetic.
OS   Plasmodium falciparum.
XX
PN   WO9832456-A1.
XX
PD   30-JUL-1998.
XX
PF   23-JAN-1998;   98WO-US01373.
XX
PR   07-FEB-1997;   97US-0037432.
PR   23-JAN-1997;   97US-0036713.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Sidney J,  Southwood S;
XX
DR   WPI; 1998-427679/36.
XX
PT   Composition containing peptide that induces cytotoxic T lymphocyte
PT   response, and helper peptide - can bind to human leucocyte antigen
PT   alleles, used to treat or prevent cancers, parasitic infections and
PT   autoimmune disease
XX
PS   Disclosure; Page 41; 51pp; English.
XX
CC   W85284-451 represent helper T-cell class II peptides, which can bind to
CC   the human leucocyte antigens (HLA) DR4w4, DR1 and DR7. The peptides
CC   are used in the course of the invention. The specification describes
CC   peptides that that induce a cytotoxic T lymphocyte (CTL) response, and
CC   T-helper peptides, that are used together to generate a CTL response for
CC   the treatment or prevention of viral, fungal, bacterial or parasitic
CC   infections (e.g. hepatitis, acquired immune deficiency syndrome or
CC   malaria) or cancer (e.g. renal or cervical carcinoma, lymphoma, prostate
CC   cancer or condyloma acuminatum). Helper T-cell peptides may be used
CC   alone to induce a helper T cell response, e.g. in cases of autoimmune
CC   disease, allograft rejection, allergy, Lyme disease, hepatitis,
CC   post-streptococcal endocarditis, glomerulonephritis and food
CC   hypersensitivity.
XX
SQ   Sequence   15 AA;
```

```
  Query Match           49.0%;  Score 25;  DB 19;  Length 15;
  Best Local Similarity  44.4%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 VMAGVGSPY 9
           |: |  :||
Db       3 vvpgaatpy 11


RESULT  29
W85376
ID   W85376 standard; peptide; 15 AA.
XX
AC   W85376;
XX
DT   16-FEB-1999  (first entry)
XX
DE   Helper T-cell class II peptide derived from SSP2 protein.
XX
KW   Helper T-cell peptide; human leucocyte antigen; HLA; DR4w4; DR1;
KW   DR7; cytotoxic T lymphocyte; CTL; hepatitis; autoimmune disease;
KW   acquired immune deficiency syndrome; malaria; cancer;
KW   allograft rejection; allergy; Lyme disease; hepatitis;
KW   post-streptococcal endocarditis; glomerulonephritis;
KW   food hypersensitivity.
XX
OS   Synthetic.
OS   Plasmodium falciparum.
XX
PN   WO9832456-A1.
XX
PD   30-JUL-1998.
XX
PF   23-JAN-1998;   98WO-US01373.
XX
PR   07-FEB-1997;   97US-0037432.
PR   23-JAN-1997;   97US-0036713.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Sidney J,  Southwood S;
XX
DR   WPI; 1998-427679/36.
XX
PT   Composition containing peptide that induces cytotoxic T lymphocyte
PT   response, and helper peptide - can bind to human leucocyte antigen
PT   alleles, used to treat or prevent cancers, parasitic infections and
PT   autoimmune disease
XX
PS   Disclosure; Page 41; 51pp; English.
XX
CC   W85284-451 represent helper T-cell class II peptides, which can bind to
CC   the human leucocyte antigens (HLA) DR4w4, DR1 and DR7. The peptides
CC   are used in the course of the invention. The specification describes
CC   peptides that that induce a cytotoxic T lymphocyte (CTL) response, and
CC   T-helper peptides, that are used together to generate a CTL response for
CC   the treatment or prevention of viral, fungal, bacterial or parasitic
CC   infections (e.g. hepatitis, acquired immune deficiency syndrome or
CC   malaria) or cancer (e.g. renal or cervical carcinoma, lymphoma, prostate
CC   cancer or condyloma acuminatum). Helper T-cell peptides may be used
CC   alone to induce a helper T cell response, e.g. in cases of autoimmune
CC   disease, allograft rejection, allergy, Lyme disease, hepatitis,
CC   post-streptococcal endocarditis, glomerulonephritis and food
CC   hypersensitivity.
XX
SQ   Sequence   15 AA;

  Query Match           49.0%;  Score 25;  DB 19;  Length 15;
  Best Local Similarity  44.4%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 VMAGVGSPY 9
           |: |  :||
Db       7 vvpgaatpy 15


RESULT  30
W85361
ID   W85361 standard; peptide; 15 AA.
XX
AC   W85361;
XX
DT   16-FEB-1999  (first entry)
XX
DE   Helper T-cell class II peptide derived from SSP2 protein.
XX
KW   Helper T-cell peptide; human leucocyte antigen; HLA; DR4w4; DR1;
KW   DR7; cytotoxic T lymphocyte; CTL; hepatitis; autoimmune disease;
KW   acquired immune deficiency syndrome; malaria; cancer;
KW   allograft rejection; allergy; Lyme disease; hepatitis;
KW   post-streptococcal endocarditis; glomerulonephritis;
KW   food hypersensitivity.
XX
OS   Synthetic.
OS   Plasmodium falciparum.
XX
PN   WO9832456-A1.
XX
PD   30-JUL-1998.
XX
PF   23-JAN-1998;   98WO-US01373.
XX
PR   07-FEB-1997;   97US-0037432.
PR   23-JAN-1997;   97US-0036713.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Sidney J,  Southwood S;
XX
DR   WPI; 1998-427679/36.
XX
PT   Composition containing peptide that induces cytotoxic T lymphocyte
PT   response, and helper peptide - can bind to human leucocyte antigen
PT   alleles, used to treat or prevent cancers, parasitic infections and
PT   autoimmune disease
XX
PS   Disclosure; Page 41; 51pp; English.
XX
CC   W85284-451 represent helper T-cell class II peptides, which can bind to
CC   the human leucocyte antigens (HLA) DR4w4, DR1 and DR7. The peptides
CC   are used in the course of the invention. The specification describes
CC   peptides that that induce a cytotoxic T lymphocyte (CTL) response, and
CC   T-helper peptides, that are used together to generate a CTL response for
CC   the treatment or prevention of viral, fungal, bacterial or parasitic
CC   infections (e.g. hepatitis, acquired immune deficiency syndrome or
CC   malaria) or cancer (e.g. renal or cervical carcinoma, lymphoma, prostate
CC   cancer or condyloma acuminatum). Helper T-cell peptides may be used
CC   alone to induce a helper T cell response, e.g. in cases of autoimmune
CC   disease, allograft rejection, allergy, Lyme disease, hepatitis,
CC   post-streptococcal endocarditis, glomerulonephritis and food
CC   hypersensitivity.
XX
SQ   Sequence   15 AA;

  Query Match           49.0%;  Score 25;  DB 19;  Length 15;
  Best Local Similarity  44.4%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 VMAGVGSPY 9
           |: |  :||
Db       5 vvpgaatpy 13


RESULT  31
```

```
W85181
ID   W85181 standard; peptide; 15 AA.
XX
AC   W85181;
XX
DT   16-FEB-1999  (first entry)
XX
DE   Helper T-cell peptide derived from a TRAP protein.
XX
KW   Helper T-cell peptide; human leucocyte antigen; HLA; DR4w4; DR1; DR7;
KW   cytotoxic T lymphocyte; CTL; hepatitis; autoimmune disease;
KW   acquired immune deficiency syndrome; malaria; cancer;
KW   allograft rejection; allergy; Lyme disease; hepatitis;
KW   post-streptococcal endocarditis; glomerulonephritis;
KW   food hypersensitivity.
XX
OS   Synthetic.
OS   Plasmodium falciparum.
XX
PN   WO9832456-A1.
XX
PD   30-JUL-1998.
XX
PF   23-JAN-1998;   98WO-US01373.
XX
PR   07-FEB-1997;   97US-0037432.
PR   23-JAN-1997;   97US-0036713.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Sidney J,  Southwood S;
XX
DR   WPI; 1998-427679/36.
XX
PT   Composition containing peptide that induces cytotoxic T lymphocyte
PT   response, and helper peptide - can bind to human leucocyte antigen
PT   alleles, used to treat or prevent cancers, parasitic infections and
PT   autoimmune disease
XX
PS   Claim 11; Page 37; 51pp; English.
XX
CC   W85138-283 represent helper T-cell peptides, which can bind to the
CC   human leucocyte antigens (HLA) DR4w4, DR1 and DR7. The peptides
CC   are used in the course of the invention. The specification describes
CC   peptides that that induce a cytotoxic T lymphocyte (CTL) response, and
CC   T-helper peptides, that are used together to generate a CTL response for
CC   the treatment or prevention of viral, fungal, bacterial or parasitic
CC   infections (e.g. hepatitis, acquired immune deficiency syndrome or
CC   malaria) or cancer (e.g. renal or cervical carcinoma, lymphoma, prostate
CC   cancer or condyloma acuminatum). Helper T-cell peptides may be used
CC   alone to induce a helper T cell response, e.g. in cases of autoimmune
CC   disease, allograft rejection, allergy, Lyme disease, hepatitis,
CC   post-streptococcal endocarditis, glomerulonephritis and food
CC   hypersensitivity.
XX
SQ   Sequence   15 AA;

  Query Match             49.0%;  Score 25;  DB 19;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 VMAGVGSPY 9
           |: |  :||
Db       4 vvpgaatpy 12


RESULT  32
W85164
ID   W85164 standard; peptide; 15 AA.
XX
AC   W85164;
XX
DT   16-FEB-1999  (first entry)
XX
DE   Helper T-cell peptide derived from a TRAP protein.
XX
KW   Helper T-cell peptide; human leucocyte antigen; HLA; DR4w4; DR1; DR7;
KW   cytotoxic T lymphocyte; CTL; hepatitis; autoimmune disease;
KW   acquired immune deficiency syndrome; malaria; cancer;
KW   allograft rejection; allergy; Lyme disease; hepatitis;
KW   post-streptococcal endocarditis; glomerulonephritis;
KW   food hypersensitivity.
XX
OS   Synthetic.
OS   Plasmodium falciparum.
XX
PN   WO9832456-A1.
XX
PD   30-JUL-1998.
XX
PF   23-JAN-1998;   98WO-US01373.
XX
PR   07-FEB-1997;   97US-0037432.
PR   23-JAN-1997;   97US-0036713.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Sidney J,  Southwood S;
XX
DR   WPI; 1998-427679/36.
XX
PT   Composition containing peptide that induces cytotoxic T lymphocyte
PT   response, and helper peptide - can bind to human leucocyte antigen
PT   alleles, used to treat or prevent cancers, parasitic infections and
PT   autoimmune disease
XX
PS   Claim 11; Page 37; 51pp; English.
XX
CC   W85138-283 represent helper T-cell peptides, which can bind to the
CC   human leucocyte antigens (HLA) DR4w4, DR1 and DR7. The peptides
CC   are used in the course of the invention. The specification describes
CC   peptides that that induce a cytotoxic T lymphocyte (CTL) response, and
CC   T-helper peptides, that are used together to generate a CTL response for
CC   the treatment or prevention of viral, fungal, bacterial or parasitic
CC   infections (e.g. hepatitis, acquired immune deficiency syndrome or
CC   malaria) or cancer (e.g. renal or cervical carcinoma, lymphoma, prostate
CC   cancer or condyloma acuminatum). Helper T-cell peptides may be used
CC   alone to induce a helper T cell response, e.g. in cases of autoimmune
CC   disease, allograft rejection, allergy, Lyme disease, hepatitis,
CC   post-streptococcal endocarditis, glomerulonephritis and food
CC   hypersensitivity.
XX
SQ   Sequence   15 AA;

  Query Match             49.0%;  Score 25;  DB 19;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 VMAGVGSPY 9
           |: |  :||
Db       3 vvpgaatpy 11


RESULT  33
W85155
ID   W85155 standard; peptide; 15 AA.
XX
AC   W85155;
XX
DT   16-FEB-1999  (first entry)
XX
DE   Helper T-cell peptide derived from a TRAP protein.
XX
KW   Helper T-cell peptide; human leucocyte antigen; HLA; DR4w4; DR1; DR7;
```

```
KW   cytotoxic T lymphocyte; CTL; hepatitis; autoimmune disease;
KW   acquired immune deficiency syndrome; malaria; cancer;
KW   allograft rejection; allergy; Lyme disease; hepatitis;
KW   post-streptococcal endocarditis; glomerulonephritis;
KW   food hypersensitivity.
XX
OS   Synthetic.
OS   Plasmodium falciparum.
XX
PN   WO9832456-A1.
XX
PD   30-JUL-1998.
XX
PF   23-JAN-1998;   98WO-US01373.
XX
PR   07-FEB-1997;   97US-0037432.
PR   23-JAN-1997;   97US-0036713.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Sidney J,  Southwood S;
XX
DR   WPI; 1998-427679/36.
XX
PT   Composition containing peptide that induces cytotoxic T lymphocyte
PT   response, and helper peptide - can bind to human leucocyte antigen
PT   alleles, used to treat or prevent cancers, parasitic infections and
PT   autoimmune disease
XX
PS   Claim 11; Page 37; 51pp; English.
XX
CC   W85138-283 represent helper T-cell peptides, which can bind to the
CC   human leucocyte antigens (HLA) DR4w4, DR1 and DR7. The peptides
CC   are used in the course of the invention. The specification describes
CC   peptides that that induce a cytotoxic T lymphocyte (CTL) response, and
CC   T-helper peptides, that are used together to generate a CTL response for
CC   the treatment or prevention of viral, fungal, bacterial or parasitic
CC   infections (e.g. hepatitis, acquired immune deficiency syndrome or
CC   malaria) or cancer (e.g. renal or cervical carcinoma, lymphoma, prostate
CC   cancer or condyloma acuminatum). Helper T-cell peptides may be used
CC   alone to induce a helper T cell response, e.g. in cases of autoimmune
CC   disease, allograft rejection, allergy, Lyme disease, hepatitis,
CC   post-streptococcal endocarditis, glomerulonephritis and food
CC   hypersensitivity.
XX
SQ   Sequence   15 AA;


  Query Match             49.0%;  Score 25;  DB 19;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 VMAGVGSPY 9
            |: |  :||
Db       5 vvpgaatpy 13


RESULT  34
Y55981
ID   Y55981 standard; Peptide; 15 AA.
XX
AC   Y55981;
XX
DT   18-FEB-2000  (first entry)
XX
DE   Human SULU1-derived peptide #3.
XX
KW   Antirheumatic; antiathritic; antiinflammatory; antiallergic; osteopathic;
KW   antipsoriatic; antiarteriosclerotic; antiasthmatic; immunosupressive;
KW   neuroprotective; cardiant; cerebroprotective; cytostatic; antidiabetic;
KW   vulnery; STE20; protein kinase; STLK2; STLK3; STLK4; STLK5, STLK6, STLK7;
KW   ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4; PAK5; antagonist;
KW   antibody; gene therapy; rheumatoid arthritis; artherosclerosis; asthma;
KW   inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;
KW   rhinitis; autoimmunity; organ transplantation; multiple sclerosis;
KW   myocardial infarction; cardiovascular disease; stroke; renal failure;
KW   oxidative stress-related neurodegenerative disorder; Parkinson's disease;
KW   amylotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;
KW   ischemic disorder; inflammation; diabetes mellitus; fibrosis; mitosis;
KW   mesangial disorder; growth regulation; wound healing; T cell activation;
KW   immunosuppressant.
XX
OS   Homo sapiens.
XX
PN   WO9953036-A2.
XX
PD   21-OCT-1999.
XX
PF   13-APR-1999;   99WO-US08150.
XX
PR   14-APR-1998;   98US-0081784.
XX
PA   (SUGE-) SUGEN INC.
XX
PI   Plowman G,  Martinez R,  Whyte D;
XX
DR   WPI; 1999-611301/52.
XX
PT   Novel kinase-related polypeptides used for the diagnosis and treatment
PT   of kinase-related diseases and disorders  -
XX
PS   Disclosure; Page 382; 387pp; English.
XX
CC   This sequence represents a peptide fragment from a novel STE20-related
CC   protein kinases. The invention relates to nucleic acid molecules encoding
CC   a kinase polypeptide selected from STLK2, STLK3, STLK4, STLK5, STLK6,
CC   STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4 and PAK5. The
CC   proteins are used to identify agonists and antagonists, and to raise
CC   antibodies. The polynucleotides are useful in gene therapy protocols. The
CC   polynucleotides, polypeptides, antibodies, antagonists and agonists may
CC   be used to treat diseases such as immune-related disorders and diseases
CC   (e.g. rheumatoid arthritis, artherosclerosis, chronic inflammatory bowel
CC   disease (e.g. Crohn's disease), asthma, osteoarthritis, psoriasis,
CC   atherosclerosis, rhinitis, autoimmunity, and organ transplantation,
CC   chronic inflammatory pelvic disease, multiple sclerosis, organ
CC   transplantation, myocardial infarction, cardiovascular disease, stroke,
CC   renal failure, oxidative stress-related neurodegenerative disorders (e.g.
CC   amylotrophic lateral sclerosis, Parkinson's disease and Leigh syndrome),
CC   cancer, cardiomyopathies, ischemic disorders, inflammatory disorders,
CC   diabetes mellitus, fibrotic and mesangial disorders. The proteins may
CC   also be useful for cell growth regulation (e.g. in wound healing), T cell
CC   activation, mitosis control, and as immunosuppressants.
XX
SQ   Sequence   15 AA;


  Query Match             49.0%;  Score 25;  DB 20;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 VGSPY 9
            ||:||
Db       6 vgtpy 10


RESULT  35
Y88496
ID   Y88496 standard; peptide; 15 AA.
XX
AC   Y88496;
XX
DT   07-AUG-2000  (first entry)
XX
DE   Peptide #2 used in identification of HLA II binding antigen epitopes.
XX
KW   Human leucocyte antigen; HLA class II; antigen epitope; pharmaceutical;
```

KW immune response; chronic viral disease; cancer; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; myasthenia gravis; AIDS;
KW allograft rejection; allergy; lyme disease; hepatitis; prostate cancer;
KW glomerulonephritis; food hypersensitivity; malaria.
XX
OS Unidentified.
XX
PN WO9961916-A1.
XX
PD 02-DEC-1999.
XX
PF 28-MAY-1999; 99WO-US12066.
XX
PR 29-MAY-1998; 98US-0087192.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Southwood S, Sidney J;
XX
DR WPI; 2000-097143/08.
XX
PT New compositions containing immunogenic peptide epitopes for various
PT HLA class II DR molecules useful for inducing helper T cell response
XX
PS Examples; Page 38; 60pp; English.
XX
CC The present invention relates to a new pharmaceutical composition
CC comprising a unit dose form of a peptide, or analogue, comprising an
CC epitope selected from those represented by peptides Y98812-Y99339 which
CC are derived from various antigens for various human leucocyte antigen
CC class DR molecules, representative of the world wide population. The
CC peptide/analogue binds to an HLA class II molecule at an IC-50 of less
CC than or equal to 1,000 nM. The present sequence is used in the
CC identification of the peptides used in the pharmaceutical composition of
CC the invention. II binding assay in the examples of the invention. The
CC pharmaceutical can be used to induce a helper T cell response. The
CC pharmaceutical focuses the immune response towards selected determinants
CC and could therefore be used in cases of chronic viral diseases and
CC cancer. Examples of diseases that can be treated using the peptide
CC containing pharmaceutical include autoimmune diseases (rheumatoid
CC arthritis, multiple sclerosis, and myasthenia gravis), allograft
CC rejection, allergies, lyme disease, hepatitis, post-streptococcal
CC endocarditis or glomerulonephritis and food hypersensitivities. The
CC peptide epitopes can be used to enhance immune responses against other
CC immunogens administered with the peptides. Diseases which can be treated
CC using immunogenic mixtures include prostate cancer, hepatitis B,
CC hepatitis C, AIDS, renal carcinoma, cervical carcinoma, lymphoma, and
CC condyloma acuminatum. The peptides may also be used to make monoclonal
CC antibodies useful as potential diagnostic or therapeutic agents. The
CC peptides may also be useful as diagnostic reagents, for example, to
CC determine the susceptibility of an individual to a treatment regimen.
CC Also, the peptides may be used to predict which individuals will be at
CC substantial risk of developing chronic infection. The selection of
CC appropriate T and B cell epitopes should allow the development of epitope
CC based vaccines particularly towards conserved epitopes of pathogens which
CC are characterized by high sequence variability such as HIV, HCV and
CC Malaria.
XX
SQ Sequence 15 AA;

Query Match 49.0%; Score 25; DB 21; Length 15;
Best Local Similarity 44.4%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 VMAGVGSPY 9
            |: |  :||
Db       7 vvpgaatpy 15
```

Search completed: December 12, 2000, 02:39:42
Job time: 12732 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: December 12, 2000, 01:15:34 ; Search time 13.09 Seconds
(without alignments)
48.480 Million cell updates/sec

Title: US-08-860-232-12
Perfect score: 51
Sequence: 1 VMAGVGSPYV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	25	49.0	19	2	I49039	T-cell receptor be
2	24	47.1	9	2	PH1591	Ig H chain V-D-J r
3	24	47.1	13	2	S47361	T-cell antigen rec
4	24	47.1	15	2	PH1610	Ig H chain V-D-J r
5	24	47.1	17	2	S33609	extensin - maize (
6	22	43.1	15	2	PA0079	malate dehydrogena
7	22	43.1	18	2	A49404	T-cell receptor be
8	21	41.2	11	2	PH0929	T-cell receptor be
9	21	41.2	13	2	S33800	chaperone, TCP1-re
10	21	41.2	18	2	A24749	neuropeptide A - b
11	20	39.2	11	2	PU0029	33K protein 3218 -
12	20	39.2	12	2	PQ0786	NADH dehydrogenase
13	20	39.2	12	2	S29830	dimethylaniline mo
14	20	39.2	14	2	PH0755	T-cell receptor be
15	20	39.2	15	2	A22789	platelet-derived g
16	20	39.2	15	2	I67525	CD33 antigen homol
17	20	39.2	16	2	A41170	photosystem II 6.1
18	20	39.2	16	2	H35141	T-cell receptor de
19	20	39.2	18	2	PQ0680	photosystem I 5.6K
20	20	39.2	18	2	I49408	cytochrome-c oxida
21	20	39.2	20	2	S46479	retinoid-X-recepto
22	19.5	38.2	20	2	A37988	acid proteinase he
23	19	37.3	8	2	E47393	neuropeptide calla
24	19	37.3	11	2	F58501	43.5K bile stone p
25	19	37.3	11	2	PT0250	Ig heavy chain CRD
26	19	37.3	13	2	S20578	ribosomal protein
27	19	37.3	14	1	NYPG14	hypothalamic tetra
28	19	37.3	14	2	PQ0698	unidentified 6.0/6
29	19	37.3	14	2	PH1766	T cell receptor al
30	19	37.3	14	2	PH1608	Ig H chain V-D-J r
31	19	37.3	15	2	S66443	NAD(P)+ transhydro
32	19	37.3	16	2	PT0224	Ig heavy chain CDR
33	19	37.3	16	2	PH1589	Ig H chain V-D-J r
34	19	37.3	16	2	B23692	transcription fact
35	19	37.3	18	2	PH1350	Ig heavy chain DJ
36	19	37.3	19	2	S68394	H+-transporting AT
37	18.5	36.3	20	2	S16362	opacity protein P.
38	18	35.3	7	2	T09512	NADH dehydrogenase
39	18	35.3	8	2	F60588	sperm-activating p
40	18	35.3	9	2	S65865	collagen alpha 2(V
41	18	35.3	11	2	PH1375	T antigen variant
42	18	35.3	11	2	E41476	probable antigen 5
43	18	35.3	14	2	PH1322	Ig heavy chain DJ
44	18	35.3	14	2	PH0915	T-cell receptor be
45	18	35.3	15	2	S08209	hypothetical prote
46	18	35.3	15	2	PQ0780	NADH dehydrogenase
47	18	35.3	15	2	S31219	30K protein - bovi
48	18	35.3	15	2	PH1449	T-cell receptor al
49	18	35.3	16	2	A48630	bothrojaracin - ja
50	18	35.3	16	2	A26393	annexin 36K chain
51	18	35.3	17	2	S40530	aleurone protein -
52	18	35.3	17	2	S09085	proteasome chain 4
53	18	35.3	18	2	A61121	serine proteinase
54	18	35.3	18	2	C56046	urinary tract ston
55	18	35.3	19	2	PH1304	Ig heavy chain DJ
56	18	35.3	20	2	B44920	2-halobenzoate 1,2
57	18	35.3	20	2	S10876	hypothetical prote
58	17	33.3	5	2	A41225	copper resistance
59	17	33.3	6	2	A27696	contraction-inhibi
60	17	33.3	8	2	PT0627	T-cell receptor be
61	17	33.3	9	2	PT0324	Ig heavy chain CRD
62	17	33.3	10	2	G60787	sperm-activating p
63	17	33.3	10	2	I60588	sperm-activating p
64	17	33.3	10	2	C39111	Ig heavy chain C r
65	17	33.3	11	2	D57789	gallbladder stone
66	17	33.3	12	2	B60228	Fc mu (IgM) recept
67	17	33.3	12	2	PH1587	Ig H chain V-D-J r
68	17	33.3	13	1	UNBO	neurotensin - bovi
69	17	33.3	13	2	A61067	neurotensin - comm
70	17	33.3	13	2	A28505	neurotensin-like p
71	17	33.3	13	2	A54326	glandular kallikre
72	17	33.3	13	2	S78766	ribosomal protein
73	17	33.3	13	2	PH0796	T-cell receptor al
74	17	33.3	13	2	A53608	neurotensin - guin
75	17	33.3	14	2	S47366	T-cell antigen rec

ALIGNMENTS

RESULT 1
I49039
T-cell receptor beta chain V-D-J-C region (V beta 6, J beta 2.7) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997
C;Accession: I49039
R;Rosenberg, W.M.; Moss, P.A.; Bell, J.I.
Eur. J. Immunol. 22, 541-549, 1992
A;Title: Variation in human T cell receptor V beta and J beta repertoire: analysis us
A;Reference number: A49039; MUID:92164737
A;Accession: I49039
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-19 <ROS>
A;Note: sequence extracted from NCBI backbone (NCBIP:90721)
C;Keywords: T-cell receptor

Query Match 49.0%; Score 25; DB 2; Length 19;
Best Local Similarity 57.1%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      3 AGVGSPY 9
          :| | ||
Db      8 SGAGIPY 14


RESULT   2
PH1591
Ig H chain V-D-J region (wild-type clone 142) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C;Accession: PH1591
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A;Reference number: PH1580; MUID:93301609
A;Accession: PH1591
A;Molecule type: DNA
A;Residues: 1-9 <LEV>
A;Experimental source: bone marrow pre-B lymphocyte
C;Keywords: immunoglobulin


  Query Match             47.1%;  Score 24;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 GSPY 9
          ||||
Db      4 GSPY 7


RESULT   3
S47361
T-cell antigen receptor VJ junction beta chain - human
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C;Accession: S47361
R;Lehner, P.J.
submitted to the EMBL Data Library, August 1994
A;Description: Human HLA-A0201 restricted recognition of influenza A is dominated by T c
A;Reference number: S47355
A;Accession: S47361
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-13 <LEH>
A;Cross-references: EMBL:Z35685; NID:g527459; PIDN:CAA84754.1; PID:g527460
C;Keywords: T-cell receptor


  Query Match             47.1%;  Score 24;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 GSPY 9
          ||||
Db      6 GSPY 9


RESULT   4
PH1610
Ig H chain V-D-J region (wild-type clone 337) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C;Accession: PH1610
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A;Reference number: PH1580; MUID:93301609
A;Accession: PH1610
A;Molecule type: DNA
A;Residues: 1-15 <LEV>
A;Experimental source: bone marrow pre-B lymphocyte
C;Keywords: immunoglobulin


  Query Match             47.1%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 GSPY 9
          ||||
Db      6 GSPY 9


RESULT   5
S33609
extensin - maize (fragment)
C;Species: Zea mays (maize)
C;Date: 19-Mar-1997 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C;Accession: S33609
R;Murphy, J.M.; Hood, E.E.
Plant Mol. Biol. 21, 885-893, 1993
A;Title: Molecular basis for extensin size heterogeneity in two maize varieties.
A;Reference number: S33609; MUID:93222485
A;Accession: S33609
A;Molecule type: protein
A;Residues: 1-17 <MUR>
C;Keywords: glycoprotein; hydroxyproline


  Query Match             47.1%;  Score 24;  DB 2;  Length 17;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GVGSPY 9
          | | ||.
Db      5 GYGPPY 10


RESULT   6
PA0079
malate dehydrogenase (EC 1.1.1.37) II - fungus (Fusarium sporotrichioides) (fragment)
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 03-Mar-1995
C;Accession: PA0079
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Reference number: PA0051
A;Accession: PA0079
A;Molecule type: protein
A;Residues: 1-15 <CHO>
C;Keywords: oxidoreductase


  Query Match             43.1%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   60.0%;  Pred. No. 4.7e+02;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 GVGSP 8
          |:| |
Db     11 GIGQP 15


RESULT   7
A49404
T-cell receptor beta chain VDJ region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C;Accession: A49404
R;Brooks, E.G.; Balk, S.P.; Aupeix, K.; Colonna, M.; Strominger, J.L.; Groh-Spies, V.
Proc. Natl. Acad. Sci. U.S.A. 90, 11787-11791, 1993
A;Title: Human T-cell receptor (TCR) alpha/beta + CD4-CD8- T cells express oligoclona
A;Reference number: A49404; MUID:94089717
A;Accession: A49404
A;Status: preliminary
A;Molecule type: mRNA
```

```
A;Residues: 1-18 <BRO>
A;Cross-references: GB:S67426; NID:g455866; PIDN:AAB29274.1; PID:g455867
A;Experimental source: alpha/beta + CD4-CD8- T cells
A;Note: sequence extracted from NCBI backbone (NCBIN:141022, NCBIP:141023)
C;Keywords: T-cell receptor


  Query Match             43.1%;  Score 22;  DB 2;  Length 18;
  Best Local Similarity   80.0%;  Pred. No. 5.7e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 MAGVG 6
            :||||
Db       5 LAGVG 9


RESULT   8
PH0929
T-cell receptor beta chain V-D-J region (clone 15) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
C;Accession: PH0929
R;Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A;Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergi
A;Reference number: PH0891; MUID:92078857
A;Accession: PH0929
A;Molecule type: mRNA
A;Residues: 1-11 <GOL>
A;Experimental source: concanavalin A-activated lymphoblast
C;Keywords: T-cell receptor


  Query Match             41.2%;  Score 21;  DB 2;  Length 11;
  Best Local Similarity   75.0%;  Pred. No. 5.3e+02;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 GSPY 9
            |:||
Db       6 GTPY 9


RESULT   9
S33800
chaperone, TCP1-related - oat
C;Species: Avena sativa (oat)
C;Date: 02-Dec-1993 #sequence_revision 27-Feb-1997 #text_change 17-Mar-1999
C;Accession: S33800
R;Mummert, E.; Grimm, R.; Speth, V.; Eckerskorn, C.; Schiltz, E.; Gatenby, A.A.; Schaefe
Nature 363, 644-648, 1993
A;Title: A TCP1-related molecular chaperone from plants refolds phytochrome to its photo
A;Reference number: S33800; MUID:93288140
A;Accession: S33800
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-13 <MUM>


  Query Match             41.2%;  Score 21;  DB 2;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 6.3e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 MAGVGSP 8
            | | |:|
Db       4 MDGPGNP 10


RESULT  10
A24749
neuropeptide A - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 28-Jul-1987 #sequence_revision 28-Jul-1987 #text_change 31-Dec-1993
C;Accession: A24749
R;Yang, H.Y.T.; Fratta, W.; Majane, E.A.; Costa, E.
Proc. Natl. Acad. Sci. U.S.A. 82, 7757-7761, 1985
A;Title: Isolation, sequencing, synthesis, and pharmacological characterization of tw
A;Reference number: A94074; MUID:86067985
A;Accession: A24749
A;Molecule type: protein
A;Residues: 1-18 <YAN>
C;Comment: The source of this peptide was brain.
C;Keywords: neuropeptide


  Query Match             41.2%;  Score 21;  DB 2;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 8.8e+02;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 GVGSPY 9
            |: ||:
Db       4 GLSSPF 9


RESULT  11
PU0029
33K protein 3218 - rice (strain Nohonbare) (fragment)
C;Species: Oryza sativa (rice)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 11-Apr-1995
C;Accession: PU0029
R;Tsugita, A.; Miyatake, N.
submitted to JIPID, April 1993
A;Reference number: PS0208
A;Accession: PU0029
A;Molecule type: protein
A;Residues: 1-11 <TSU>
A;Experimental source: bran
C;Comment: molecular weight 33K, pI 6.0.


  Query Match             39.2%;  Score 20;  DB 2;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+02;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GVGSPY 9
            | | |:
Db       2 GEGGPF 7


RESULT  12
PQ0786
NADH dehydrogenase (EC 1.6.99.3) 26K chain - fava bean mitochondrion (fragment)
N;Alternate names: complex I 26K chain; NADH--ubiquinone reductase 26K chain
C;Species: mitochondrion Vicia faba (fava bean)
C;Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 17-Mar-1999
C;Accession: PQ0786
R;Leterme, S.; Boutry, M.
Plant Physiol. 102, 435-443, 1993
A;Title: Purification and preliminary characterization of mitochondrial complex I (NA
A;Reference number: PQ0775; MUID:94151437
A;Accession: PQ0786
A;Molecule type: protein
A;Residues: 1-12 <LET>
C;Comment: Complex I, mitochondrial NADH-ubiquiquinone reductase, is the first of the
 ranging from 5K to 75K.
C;Comment: This enzyme catalyzes electron transfer from endogenous NADH to ubiquinone
C;Genetics:
A;Genome: mitochondrion
C;Keywords: electron transfer; mitochondrion; oxidoreductase


  Query Match             39.2%;  Score 20;  DB 2;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 9e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GVGSPY 9
            ||  ||
```

```
Db      3 GVPIPY 8


RESULT  13
S29830
dimethylaniline monooxygenase (N-oxide-forming) (EC 1.14.13.8), hepatic - crab-eating ma
N;Alternate names: flavin-containing monooxygenase
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 19-Mar-1997 #sequence_revision 24-Mar-1999 #text_change 07-May-1999
C;Accession: S29830
R;Sadeque, A.J.M.; Thummel, K.E.; Rettie, A.E.
Biochim. Biophys. Acta 1162, 127-134, 1993
A;Title: Purification of macacque liver flavin-containing monooxygenase: A form of the e
A;Reference number: S29830; MUID:93192283
A;Accession: S29830
A;Molecule type: protein
A;Residues: 1-12 <SAD>
A;Experimental source: liver
C;Keywords: FAD; flavoprotein; microsome; monooxygenase; NADP; oxidoreductase
F;1-12/Region: beta-alpha-beta FAD nucleotide-binding fold (fragment)


  Query Match             39.2%;  Score 20;  DB 2;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 AGVG 6
          ||||
Db      8 AGVG 11


RESULT  14
PH0755
T-cell receptor beta chain (QA11.3.2) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
C;Accession: PH0755
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846
A;Accession: PH0755
A;Molecule type: mRNA
A;Residues: 1-14 <CAS>
A;Cross-references: EMBL:X60849; NID:g53876; PIDN:CAA43240.1; PID:g53877
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor


  Query Match             39.2%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   75.0%;  Pred. No. 1.1e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 GSPY 9
          | ||
Db      6 GQPY 9


RESULT  15
A22789
platelet-derived growth factor chain B - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 31-Mar-1988 #sequence_revision 02-Jun-1988 #text_change 18-Jun-1993
C;Accession: A22789
R;Stroobant, P.; Waterfield, M.D.
EMBO J. 12, 2963-2967, 1984
A;Title: Purification and properties of porcine platelet-derived growth factor.
A;Reference number: A22789
A;Accession: A22789
A;Molecule type: protein
A;Residues: 1-15 <STR>
C;Superfamily: platelet-derived growth factor
C;Keywords: growth factor; mitogen


  Query Match             39.2%;  Score 20;  DB 2;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 VGSPYV 10
          :||| |
Db      2 LGSPAV 7


RESULT  16
I67525
CD33 antigen homolog - mouse (fragment)
C;Species: Mus sp. (mouse)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 05-Jun-1998
C;Accession: I67525
R;Chies, J.A.; Lembezat, M.P.; Freitas, A.A.
Eur. J. Immunol. 24, 1657-1664, 1994
A;Title: Entry of B lymphocytes into the persistent cell pool in non-immunized mice i
A;Reference number: I53392; MUID:94298870
A;Accession: I67525
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-15 <RES>
A;Cross-references: GB:S71349; NID:g550037
C;Genetics:
A;Gene: Ig VH7183


  Query Match             39.2%;  Score 20;  DB 2;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 GSPYV 10
          || ||
Db      9 GSSYV 13


RESULT  17
A41170
photosystem II 6.1K protein - Chlamydomonas reinhardtii (fragment)
C;Species: Chlamydomonas reinhardtii
C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 18-Jun-1993
C;Accession: A41170
R;de Vitry, C.; Diner, B.A.; Popot, J.L.
J. Biol. Chem. 266, 16614-16621, 1991
A;Title: Photosystem II particles from Chlamydomonas reinhardtii. Purification, molec
A;Reference number: A41170; MUID:91358452
A;Accession: A41170
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-16 <DEV>


  Query Match             39.2%;  Score 20;  DB 2;  Length 16;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+03;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GVGSPY 9
          | | |:
Db     10 GTGLPF 15


RESULT  18
H35141
T-cell receptor delta chain V region (105.23) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 31-Aug-1990 #sequence_revision 31-Aug-1990 #text_change 30-May-1997
C;Accession: H35141
R;Sim, G.K.; Augustin, A.
Cell 61, 397-405, 1990
```

```
A;Title: Dominantly inherited expression of BID, an invariant undiversified T cell recep
A;Reference number: A35141; MUID:90242386
A;Accession: H35141
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-16 <SIM>
C;Keywords: T-cell receptor


  Query Match             39.2%;  Score 20;  DB 2;  Length 16;
  Best Local Similarity   25.0%;  Pred. No. 1.2e+03;
  Matches    2;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 VMAGVGSP 8
          :: |: :|
Db      6 ILGGIRAP 13


RESULT  19
PQ0680
photosystem I 5.6K K chain - common tobacco (fragment)
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 17-Mar-1999
C;Accession: PQ0680
R;Obokata, J.; Mikami, K.; Hayashida, N.; Nakamura, M.; Sugiura, M.
Plant Physiol. 102, 1259-1267, 1993
A;Title: Molecular heterogeneity of photosystem I. psaD, psaE, psaF, psaH and psaL are a
A;Reference number: PQ0667; MUID:94105345
A;Accession: PQ0680
A;Molecule type: protein
A;Residues: 1-18 <OBO>
C;Keywords: chloroplast; photosynthesis; photosystem I; thylakoid


  Query Match             39.2%;  Score 20;  DB 2;  Length 18;
  Best Local Similarity   75.0%;  Pred. No. 1.4e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 VGSP 8
          :|||
Db      3 IGSP 6


RESULT  20
I49408
cytochrome-c oxidase (EC 1.9.3.1) chain Va - western wild mouse (fragment)
C;Species: Mus spretus (western wild mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I49408
R;Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.H
Mamm. Genome 5, 349-355, 1994
A;Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
A;Reference number: I48934; MUID:94319082
A;Accession: I49408
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-18 <RES>
A;Cross-references: EMBL:U05699; NID:g497020; PIDN:AAB60466.1; PID:g497021
C;Keywords: electron transfer; membrane-associated complex; oxidoreductase; respiratory


  Query Match             39.2%;  Score 20;  DB 2;  Length 18;
  Best Local Similarity   60.0%;  Pred. No. 1.4e+03;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 GVGSP 8
          |: ||
Db      6 GISSP 10


RESULT  21
S46479
retinoid-X-receptor-gamma - chicken
C;Species: Gallus gallus (chicken)
C;Date: 15-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 28-May-1999
C;Accession: S46479
R;Seleiro, E.A.P.; Darling, D.; Brickell, P.M.
Biochem. J. 301, 283-288, 1994
A;Title: The chicken retinoid-X-receptor-gamma gene gives rise to two distinct specie
A;Reference number: S46478; MUID:94311845
A;Accession: S46479
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-20 <SEL>
A;Cross-references: GB:S72435; NID:g619294; PIDN:AAB31348.1; PID:g619295


  Query Match             39.2%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 GVGSPY 9
          |: :||
Db      4 GMQAPY 9


RESULT  22
A37988
acid proteinase heavy chain - slime mold (Physarum polycephalum) (fragment)
C;Species: Physarum polycephalum
C;Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 30-Sep-1993
C;Accession: A37988
R;Murakami-Murofushi, K.; Takahashi, T.; Minowa, Y.; Iino, S.; Takeuchi, T.; Kitagaki
J. Biol. Chem. 265, 19898-19903, 1990
A;Title: Purification and characterization of a novel intracellular acid proteinase f
A;Reference number: A37988; MUID:91060608
A;Accession: A37988
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <MUR>


  Query Match             38.2%;  Score 19.5;  DB 2;  Length 20;
  Best Local Similarity   54.5%;  Pred. No. 1.9e+03;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    3;  Gaps    1;

Qy      3 AGVGS---PYV 10
          |||     |||
Db      1 AGVDGYIVPYV 11


RESULT  23
E47393
neuropeptide callatostatin 5 - bluebottle fly (Calliphora vomitoria)
C;Species: Calliphora vomitoria
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C;Accession: E47393
R;Duve, H.; Johnsen, A.H.; Scott, A.G.; Yu, C.G.; Yagi, K.J.; Tobe, S.S.; Thorpe, A.
Proc. Natl. Acad. Sci. U.S.A. 90, 2456-2460, 1993
A;Title: Callatostatins: neuropeptides from the blowfly Calliphora vomitoria with seq
A;Reference number: A47393; MUID:93211980
A;Accession: E47393
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-8 <DUV>
A;Experimental source: whole flies
A;Note: sequence extracted from NCBI backbone (NCBIP:128482)


  Query Match             37.3%;  Score 19;  DB 2;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 1.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 GSPY 9
          | ||
Db      1 GPPY 4
```

```
RESULT  24
F58501
43.5K bile stone protein - unidentified bacterium (fragment)
C;Species: unidentified bacterium
C;Date: 07-Feb-1997 #sequence_revision 07-Feb-1997 #text_change 10-Jul-1998
C;Accession: F58501
R;Binette, J.P.; Binette, M.B.
submitted to the Protein Sequence Database, October 1996
A;Description: The proteins of kidney and gallbladder stones.
A;Reference number: A58501
A;Accession: F58501
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-11 <BIN>
A;Experimental source: human bile with stones
A;Note: 6-Asn and 8-Ala were also found


  Query Match             37.3%;  Score 19;  DB 2;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 VMAGVGSP 8
            |  ||  |
Db       2 VKIGVAGP 9


RESULT  25
PT0250
Ig heavy chain CRD3 region (clone 2-109B) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0250
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337
A;Accession: PT0250
A;Molecule type: DNA
A;Residues: 1-11 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin


  Query Match             37.3%;  Score 19;  DB 2;  Length 11;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 GSPY 9
            | ||
Db       7 GRPY 10


RESULT  26
S20578
ribosomal protein L36 - Cryptomonas sp. chloroplast (fragment)
C;Species: chloroplast Cryptomonas sp.
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Sep-1997
C;Accession: S20578
R;Douglas, S.E.
FEBS Lett. 298, 93-96, 1992
A;Title: A secY homologue is found in the plastid genome of Cryptomonas Phi.
A;Reference number: S20577; MUID:92183838
A;Accession: S20578
A;Molecule type: DNA
A;Residues: 1-13 <DOU>
A;Cross-references: EMBL:X62348; NID:g11300; PID:g11302
A;Note: the source is designated as Cryptomonas phi
C;Genetics:
A;Gene: rpl36
A;Genome: chloroplast
C;Keywords: chloroplast


  Query Match             37.3%;  Score 19;  DB 2;  Length 13;
  Best Local Similarity   42.9%;  Pred. No. 1.5e+03;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 VMAGVGS 7
            |:: :||
Db       3 VVSSIGS 9


RESULT  27
NYPG14
hypothalamic tetradecapeptide - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 23-Aug-1996
C;Accession: A01419
R;Schlesinger, D.H.; Niall, H.D.; Linthicum, G.L.; Dupont, A.; Schally, A.V.
submitted to the Atlas, November 1976
A;Reference number: A01419
A;Accession: A01419
A;Molecule type: protein
A;Residues: 1-14 <SCH>
C;Superfamily: hypothalamic tetradecapeptide
C;Keywords: amidated carboxyl end; hypothalamus
F;14/Modified site: amidated carboxyl end (Tyr) #status experimental


  Query Match             37.3%;  Score 19;  DB 1;  Length 14;
  Best Local Similarity   75.0%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 SPYV 10
            |||:
Db       6 SPYL 9


RESULT  28
PQ0698
unidentified 6.0/60K protein [imported] - rice (fragment)
C;Species: Oryza sativa (rice)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: PQ0698
R;Komatsu, S.; Kajiwara, H.; Hirano, H.
Theor. Appl. Genet. 86, 935-942, 1993
A;Title: A rice protein library; a data-file of rice proteins separated by two-dimens
A;Reference number: PQ0696
A;Accession: PQ0698
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <KOM>


  Query Match             37.3%;  Score 19;  DB 2;  Length 14;
  Best Local Similarity   60.0%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VMAGV 5
            ::|||
Db       4 ILAGV 8


RESULT  29
PH1766
T cell receptor alpha chain V region (clone 2V alpha 7.2-1) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH1766
R;Porcelli, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.
J. Exp. Med. 178, 1-16, 1993
A;Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blo
```

A;Reference number: PH1754; MUID:93301585
A;Accession: PH1766
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-14 <POR>

Query Match 37.3%; Score 19; DB 2; Length 14;
Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AGVGSPY 9
||: | |
Db 3 AGLDSNY 9

RESULT 30
PH1608
Ig H chain V-D-J region - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C;Accession: PH1608; PH1603
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A;Reference number: PH1580; MUID:93301609
A;Accession: PH1608
A;Molecule type: DNA
A;Residues: 1-14 <LEV>
A;Experimental source: bone marrow pre-B lymphocyte, wild-type clone 335
A;Accession: PH1603
A;Molecule type: DNA
A;Residues: 1-14 <LEV2>
A;Experimental source: bone marrow pre-B lymphocyte, wild-type clone 324
C;Keywords: immunoglobulin

Query Match 37.3%; Score 19; DB 2; Length 14;
Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AGVGSPY 9
| ||: |
Db 2 ARVGNSY 8

RESULT 31
S66443
NAD(P)+ transhydrogenase (B-specific) (EC 1.6.1.1) - Rhodospirillum rubrum (fragments)
N;Alternate names: proton-translocating transhydrogenase
C;Species: Rhodospirillum rubrum
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
C;Accession: S66443
R;Diggle, C.; Cotton, N.P.J.; Grimley, R.L.; Quirk, P.G.; Thomas, C.M.; Jackson, J.B.
Eur. J. Biochem. 232, 315-326, 1995
A;Title: Conformational dynamics of a mobile loop in the NAD(H)-binding subunit of proto
A;Reference number: S66443; MUID:96048062
A;Accession: S66443
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-5;6-10;11-15 <DIG>
C;Keywords: oxidoreductase

Query Match 37.3%; Score 19; DB 2; Length 15;
Best Local Similarity 42.9%; Pred. No. 1.8e+03;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AGVGSPY 9
||:| :
Db 9 AGMGEEF 15

RESULT 32
PT0224
Ig heavy chain CDR3 region (clone 1-91) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0224
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity an
A;Reference number: PT0222; MUID:91108337
A;Accession: PT0224
A;Molecule type: DNA
A;Residues: 1-16 <YAM>
A;Experimental source: B lymphocyte
A;Note: the authors translated the stop codon for residue 9 as X
C;Keywords: heterotetramer; immunoglobulin

Query Match 37.3%; Score 19; DB 2; Length 16;
Best Local Similarity 42.9%; Pred. No. 1.9e+03;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MAGVGSP 8
: | |:|
Db 10 LLGPGNP 16

RESULT 33
PH1589
Ig H chain V-D-J region (wild-type clone 140) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C;Accession: PH1589
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less m
A;Reference number: PH1580; MUID:93301609
A;Accession: PH1589
A;Molecule type: DNA
A;Residues: 1-16 <LEV>
A;Experimental source: bone marrow pre-B lymphocyte
C;Keywords: immunoglobulin

Query Match 37.3%; Score 19; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 1.9e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GSPY 9
|||:
Db 8 GSPH 11

RESULT 34
B23692
transcription factor chain A11, CCAAT-binding - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Oct-1991 #sequence_revision 04-Oct-1991 #text_change 30-Sep-1993
C;Accession: B23692
R;Vuorio, T.; Maity, S.N.; de Crombrugghe, B.
J. Biol. Chem. 265, 22480-22486, 1990
A;Title: Purification and molecular cloning of the "A" chain of a rat heteromeric CCA
A;Reference number: A23692; MUID:91093096
A;Accession: B23692
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-16 <VUO>
A;Cross-references: GB:J05701

Query Match 37.3%; Score 19; DB 2; Length 16;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy       3 AGVGSPY 9
           ||  ||:
Db       7 AGEPSPW 13


RESULT  35
PH1350
Ig heavy chain DJ region (clone C100-109R) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PH1350
R;Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A;Title: Predominance of fetal type DJH joining in young children with B precursor lymph
A;Reference number: PH1302; MUID:93094761
A;Accession: PH1350
A;Molecule type: DNA
A;Residues: 1-18 <WAS>
A;Note: the authors translated the stop codons for residues 2 and 11 as X
C;Keywords: heterotetramer; immunoglobulin


  Query Match             37.3%;  Score 19;  DB 2;  Length 18;
  Best Local Similarity   42.9%;  Pred. No. 2.1e+03;
  Matches    3;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 MAGVGSP 8
           : | |:|
Db      12 LLGPGNP 18


Search completed: December 12, 2000, 02:44:22
Job time: 5328 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: December 12, 2000, 01:21:08 ; Search time 17.49 Seconds
(without alignments)
53.387 Million cell updates/sec

Title: US-08-860-232-12
Perfect score: 51
Sequence: 1 VMAGVGSPYV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	51.0	13	2	Q9RG00	Q9rg00 mycoplasma
2	26	51.0	15	2	Q9RG03	Q9rg03 mycoplasma
3	26	51.0	19	2	Q9RFZ7	Q9rfz7 mycoplasma
4	24	47.1	15	2	Q9RQ22	Q9rq22 salmonella
5	23	45.1	17	10	Q9S8Y3	Q9s8y3 lupinus arb
6	23	45.1	20	5	Q9TWN0	Q9twn0 macrobdella
7	22	43.1	14	12	Q9PY99	Q9py99 murine hepa
8	22	43.1	17	4	Q9UCN0	Q9ucn0 homo sapien
9	21	41.2	12	10	P82325	P82325 pisum sativ
10	21	41.2	12	11	Q61331	Q61331 mus musculu
11	21	41.2	16	6	O77489	O77489 tupaia glis
12	21	41.2	18	6	Q9TQR9	Q9tqr9 mammuthus p
13	21	41.2	20	2	Q47614	Q47614 escherichia
14	21	41.2	20	4	O75318	O75318 homo sapien
15	20	39.2	14	10	P82322	P82322 pisum sativ
16	20	39.2	15	2	Q9R599	Q9r599 micrococcus
17	20	39.2	15	4	Q9UC85	Q9uc85 homo sapien
18	20	39.2	15	10	Q9S8F1	Q9s8f1 zea mays (m
19	20	39.2	15	12	Q86869	Q86869 lymphocytic
20	20	39.2	18	6	Q9TT30	Q9tt30 saguinus la
21	20	39.2	18	8	Q9T2J5	Q9t2j5 nicotiana t
22	20	39.2	18	11	Q62532	Q62532 mus spretus
23	20	39.2	18	13	Q91380	Q91380 gallus gall
24	20	39.2	19	10	Q9S8E2	Q9s8e2 spinacia ol
25	20	39.2	19	11	Q9QUY4	Q9quy4 rattus sp.
26	20	39.2	19	13	Q9PS70	Q9ps70 gallus gall
27	19	37.3	9	4	Q9UQW0	Q9uqw0 homo sapien
28	19	37.3	13	2	Q9RFZ4	Q9rfz4 mycoplasma
29	19	37.3	15	6	Q9TRG9	Q9trg9 bos taurus
30	19	37.3	15	7	Q9TNQ3	Q9tnq3 homo sapien
31	19	37.3	16	4	Q9UC53	Q9uc53 homo sapien
32	19	37.3	16	8	Q9T2Q4	Q9t2q4 brassica na
33	19	37.3	16	8	Q9T2V8	Q9t2v8 homo sapien
34	19	37.3	17	8	Q07055	Q07055 crithidia f
35	19	37.3	19	10	Q9S882	Q9s882 chlamydomon
36	18.5	36.3	16	2	Q45663	Q45663 bacillus su
37	18	35.3	7	12	Q07624	Q07624 rous sarcom
38	18	35.3	10	4	Q9UE32	Q9ue32 homo sapien
39	18	35.3	11	4	Q9UKZ3	Q9ukz3 homo sapien
40	18	35.3	12	4	Q9UCQ5	Q9ucq5 homo sapien
41	18	35.3	12	10	Q9S8F0	Q9s8f0 zea mays (m
42	18	35.3	13	13	P82064	P82064 limnodynast
43	18	35.3	15	6	Q9TRH9	Q9trh9 bos taurus
44	18	35.3	16	4	Q9UC99	Q9uc99 homo sapien
45	18	35.3	16	6	Q9TQY6	Q9tqy6 oryctolagus
46	18	35.3	16	13	P82390	P82390 litoria aur
47	18	35.3	16	13	Q9PRZ4	Q9prz4 bothrops ja
48	18	35.3	18	2	Q9R334	Q9r334 pseudomonas
49	18	35.3	19	2	Q47895	Q47895 fremyella d
50	18	35.3	19	2	Q9R4A3	Q9r4a3 escherichia
51	18	35.3	19	4	Q9UCL3	Q9ucl3 homo sapien
52	18	35.3	19	6	Q9TRP6	Q9trp6 bos taurus
53	18	35.3	19	6	Q9TR38	Q9tr38 sus scrofa
54	18	35.3	19	13	P87484	P87484 gallus gall
55	18	35.3	20	2	Q9R5Q5	Q9r5q5 burkholderi
56	18	35.3	20	2	Q9R4C0	Q9r4c0 actinoplane
57	18	35.3	20	2	Q9R5A6	Q9r5a6 thiobacillu
58	18	35.3	20	8	Q9T2N9	Q9t2n9 gallus gall
59	18	35.3	20	12	Q83272	Q83272 cucumber mo
60	18	35.3	20	13	Q9PS38	Q9ps38 rana catesb
61	17	33.3	8	2	Q9R7T2	Q9r7t2 escherichia
62	17	33.3	8	13	P82079	P82079 limnodynast
63	17	33.3	10	13	Q9PRY8	Q9pry8 triakis scy
64	17	33.3	11	11	Q9QXM6	Q9qxm6 mus musculu
65	17	33.3	12	2	P95606	P95606 alcaligenes
66	17	33.3	12	2	Q46712	Q46712 plasmid r10
67	17	33.3	15	10	Q9S8Z0	Q9s8z0 hordeum vul
68	17	33.3	15	11	Q9QVD7	Q9qvd7 rattus sp.
69	17	33.3	15	11	Q9QVC4	Q9qvc4 rattus norv
70	17	33.3	15	12	Q86867	Q86867 lymphocytic
71	17	33.3	15	13	Q9PRM3	Q9prm3 gallus gall
72	17	33.3	16	2	Q9R5E4	Q9r5e4 aeromonas h
73	17	33.3	16	3	Q02160	Q02160 pichia past
74	17	33.3	16	4	Q16033	Q16033 homo sapien
75	17	33.3	16	12	Q07625	Q07625 rous sarcom

ALIGNMENTS

RESULT 1
Q9RG00
ID Q9RG00 PRELIMINARY; PRT; 13 AA.
AC Q9RG00;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE HYPOTHETICAL 1.5 KDA PROTEIN (FRAGMENT).
OS Mycoplasma capricolum subsp. capricolum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC capricolum group.
RN [1]

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=8035;
RA   Thiaucourt F., Lorenzon S., David A.;
RT   "Phylogeny of the Mycoplasma mycoides cluster as shown by sequencing
RT   of a putative membrane protein gene.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF162995; AAF15247.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1459 MW;  0B63638AED35573B CRC64;


  Query Match             51.0%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 VGSPYV 10
          ||:||:
Db      2 VGTPYL 7


RESULT   2
Q9RG03
ID   Q9RG03      PRELIMINARY;      PRT;    15 AA.
AC   Q9RG03;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   HYPOTHETICAL 1.7 KDA PROTEIN (FRAGMENT).
OS   Mycoplasma capricolum subsp. capricolum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   capricolum group.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=7986;
RA   Thiaucourt F., Lorenzon S., David A.;
RT   "Phylogeny of the Mycoplasma mycoides cluster as shown by sequencing
RT   of a putative membrane protein gene.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF162994; AAF15244.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   15 AA;  1721 MW;  0B636A0DDD6B6DE4 CRC64;


  Query Match             51.0%;  Score 26;  DB 2;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 VGSPYV 10
          ||:||:
Db      4 VGTPYL 9


RESULT   3
Q9RFZ7
ID   Q9RFZ7      PRELIMINARY;      PRT;    19 AA.
AC   Q9RFZ7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   HYPOTHETICAL 2.1 KDA PROTEIN (FRAGMENT).
OS   Mycoplasma capricolum subsp. capricolum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   capricolum group.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CALIF KID;
RA   Thiaucourt F., Lorenzon S., David A.;
RT   "Phylogeny of the Mycoplasma mycoides cluster as shown by sequencing
RT   of a putative membrane protein gene.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF162997; AAF15250.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   19 AA;  2104 MW;  72B6F49DFD6B6DA4 CRC64;


  Query Match             51.0%;  Score 26;  DB 2;  Length 19;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 VGSPYV 10
          ||:||:
Db      8 VGAPYL 13


RESULT   4
Q9RQ22
ID   Q9RQ22      PRELIMINARY;      PRT;    15 AA.
AC   Q9RQ22;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   FERRIC HYDROXAMATE UPTAKE PROTEIN (FRAGMENT).
GN   FHUB.
OS   Salmonella typhi.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ISP1820;
RA   Morrow B.J., Graham J.E., Curtiss R. III.;
RT   "Genomic subtractive hybridization and selective capture of
RT   transcribed sequences identify a novel Salmonella typhimurium fimbrial
RT   operon and putative transcriptional regulator that are absent from the
RT   Salmonella typhi genome.";
RL   Infect. Immun. 67:5106-5116(1999).
DR   EMBL; AF134977; AAD54116.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   15 AA;  1825 MW;  036E36EB6455E616 CRC64;


  Query Match             47.1%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   60.0%;  Pred. No. 3.7e+02;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 VGSPY 9
          :|:||
Db      1 IGAPY 5


RESULT   5
Q9S8Y3
ID   Q9S8Y3      PRELIMINARY;      PRT;    17 AA.
AC   Q9S8Y3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   L-ASPARAGINASE ISOFORM A (EC 3.5.1.1) (FRAGMENT).
OS   Lupinus arboreus (Tree lupine).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Fabales; Fabaceae;
OC   Papilionoideae; Lupinus.
RN   [1]
RP   SEQUENCE.
RA   Lough T.J., Chang K.S., Carne A., Monk B.C., Reynolds P.H.,
RA   Farnden K.J.;
RL   Phytochemistry 31:1519-1527(1992).
SQ   SEQUENCE   17 AA;  1703 MW;  9AEDD9691F7F0807 CRC64;


  Query Match             45.1%;  Score 23;  DB 10;  Length 17;
  Best Local Similarity   80.0%;  Pred. No. 6.5e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      3 AGVGS 7
          ||:||
Db      4 AGIGS 8


RESULT   6
Q9TWN0
ID   Q9TWN0      PRELIMINARY;      PRT;    20 AA.
AC   Q9TWN0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   SIALIDASE L (FRAGMENT).
OS   Macrobdella decora (North American leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arynchobdellida; Hirudiniformes; Hirudinidae; Macrobdella.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 94308136.
RA   Chou M.Y., Li S.C., Kiso M., Hasegawa A., Li Y.T.;
RT   "Purification and characterization of sialidase L, a NeuAc alpha
RT   2-->3Gal-specific sialidase.";
RL   J. Biol. Chem. 269:18821-18826(1994).
SQ   SEQUENCE   20 AA;  1967 MW;  5C1026C367C28DBC CRC64;


  Query Match             45.1%;  Score 23;  DB 5;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 AGVGS 7
          ||:||
Db     10 AGIGS 14


RESULT   7
Q9PY99
ID   Q9PY99      PRELIMINARY;      PRT;    14 AA.
AC   Q9PY99;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   NON-STRUCTURAL PROTEIN.
OS   murine hepatitis virus strain 2.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Coronaviridae; Coronavirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MHV-2;
RA   Das Sarma J., Hingley S.T., Lai M.M.C., Weiss S.R., Lavi E.;
RT   "Pathogenesis and sequence analysis of mouse hepatitis virus type 2:
RT   an experimental model system of acute meningitis and hepatitis in
RT   mice.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF201929; AAF19387.1; -.
SQ   SEQUENCE   14 AA;  1534 MW;  C2FD164C12169242 CRC64;


  Query Match             43.1%;  Score 22;  DB 12;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 8.1e+02;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MAGVGSPY 9
          | |:|  |
Db      1 MLGIGLVY 8


RESULT   8
Q9UCN0
ID   Q9UCN0      PRELIMINARY;      PRT;    17 AA.
AC   Q9UCN0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   TRANSFERRIN RECEPTOR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92375195.
RA   Chicz R.M., Urban R.G., Lane W.S., Gorga J.C., Stern L.J.,
RA   Vignali D.A., Strominger J.L.;
RT   "Predominant naturally processed peptides bound to HLA-DR1 are derived
RT   from MHC-related molecules and are heterogeneous in size.";
RL   Nature 358:764-768(1992).
SQ   SEQUENCE   17 AA;  2035 MW;  A7DEDA39A2538A88 CRC64;


  Query Match             43.1%;  Score 22;  DB 4;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 9.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 SPYV 10
          ||||
Db      8 SPYV 11


RESULT   9
P82325
ID   P82325      PRELIMINARY;      PRT;    12 AA.
AC   P82325;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   UNKNOWN PROTEIN FROM 2D-PAGE OF THYLAKOID (SPOT106) (FRAGMENT).
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Fabales; Fabaceae;
OC   Papilionoideae; Pisum.
RN   [1]
RP   SEQUENCE, SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.
RC   STRAIN=CV. DE GRACE; TISSUE=LEAF;
RA   Peltier J.B., Friso G., Kalume D.E., Roepstorff P., Nilsson F.,
RA   Adamska I., van Wijk K.J.;
RT   "Proteomics of the chloroplast.";
RL   Plant Cell 12:0-0(2000).
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE LUMEN OR
CC       PERIPHERY.
CC   -!- DEVELOPMENTAL STAGE: UNFOLDED AND FULLY DEVELOPED LEAVES.
KW   Chloroplast; Thylakoid membrane.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1236 MW;  CEAC7ADC02633452 CRC64;


  Query Match             41.2%;  Score 21;  DB 10;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 AGVGSP 8
          |||  |
Db      3 AGVNKP 8


RESULT  10
Q61331
ID   Q61331      PRELIMINARY;      PRT;    12 AA.
AC   Q61331;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   N-ACETYLGLUCOSAMINE GALACTOSYLTRANSFERASE (BETA1-4GT) (FRAGMENT).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 89033997.
RA   Nakazawa K., Ando T., Kimura T., Narimatsu H.;
RT   "Cloning and sequencing of a full-length cDNA of mouse N-
RT   acetylglucosamine (beta 1-4)galactosyltransferase.";
RL   J. Biochem. 104:165-168(1988).
DR   EMBL; D00315; BAA00217.1; -.
KW   Transferase; Glycosyltransferase.
FT   NON_TER       1      1
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1283 MW;  304EA40668387728 CRC64;


  Query Match             41.2%;  Score 21;  DB 11;  Length 12;
  Best Local Similarity   37.5%;  Pred. No. 1e+03;
  Matches    3;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 AGVGSPYV 10
          :|:|  |:
Db      4 SGLGKTYL 11


RESULT  11
O77489
ID   O77489      PRELIMINARY;      PRT;    16 AA.
AC   O77489;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   D4 DOPAMINE RECEPTOR (D4DR) (FRAGMENT).
OS   Tupaia glis (Tree shrew).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaia.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Inoue-Murayama M., Takenaka O., Murayama Y.;
RT   "Origin and divergence of tandem repeats of primate D4 dopamine
RT   receptor genes.";
RL   Primates 39:217-224(1998).
DR   EMBL; AB016198; BAA32036.1; -.
FT   NON_TER       1      1
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1577 MW;  3865AEE77FB63E09 CRC64;


  Query Match             41.2%;  Score 21;  DB 6;  Length 16;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 GVGSP 8
          | |||
Db      2 GPGSP 6


RESULT  12
Q9TQR9
ID   Q9TQR9      PRELIMINARY;      PRT;    18 AA.
AC   Q9TQR9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   VON WILLEBRAND FACTOR (FRAGMENT).
GN   VWF.
OS   Mammuthus primigenius (Siberian woolly mammoth).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Proboscidea; Elephantidae; Mammuthus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 20022977.
RA   Greenwood A.D., Capelli C., Possnert G., Paabo S.;
RT   "Nuclear DNA sequences from late Pleistocene megafauna.";
RL   Mol. Biol. Evol. 16:1466-1473(1999).
DR   EMBL; AF154874; AAF12750.1; -.
DR   EMBL; AF154873; AAF12749.1; -.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  1914 MW;  DFCB484B41F69236 CRC64;


  Query Match             41.2%;  Score 21;  DB 6;  Length 18;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 VGSPYV 10
          | :|||
Db      5 VTTPYV 10


RESULT  13
Q47614
ID   Q47614      PRELIMINARY;      PRT;    20 AA.
AC   Q47614;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   URF WITH HOMOLOGY TO RRNE URF2.
GN   URF.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12, HFRC;
RX   MEDLINE; 91232952.
RA   Albrechtsen B., Ross B.M., Squires C., Squires C.L.;
RT   "Transcriptional termination sequence at the end of the Escherichia
RT   coli ribosomal RNA G operon: complex terminators and
RT   antitermination.";
RL   Nucleic Acids Res. 19:1845-1852(1991).
DR   EMBL; X56780; CAA40098.1; -.
SQ   SEQUENCE   20 AA;  2162 MW;  D952ACD71417E163 CRC64;


  Query Match             41.2%;  Score 21;  DB 2;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 VGSPY 9
          :| ||
Db      6 LGKPY 10


RESULT  14
O75318
ID   O75318      PRELIMINARY;      PRT;    20 AA.
AC   O75318;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   UBIQUITIN HYDROLYZING ENZYME 1 (FRAGMENT).
GN   UBH1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 98277453.
RA   Hansen-Hagge T.E., Janssen J.W., Hameister H., Papa F.R., Zechner U.,
RA   Seriu T., Jauch A., Becke D., Hochstrasser M., Bartram C.R.;
RT   "An evolutionarily conserved gene on human chromosome 5q33-q34, UBH1,
RT   encodes a novel deubiquitinating enzyme.";
RL   Genomics 49:411-418(1998).
DR   EMBL; AF022793; AAC23451.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   20 AA;  2214 MW;  DB9C921FD3802D22 CRC64;
```

```
  Query Match             41.2%;  Score 21;  DB 4;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 GSPYV 10
          |:|:|
Db      7 GNPFV 11


RESULT  15
P82322
ID   P82322      PRELIMINARY;      PRT;    14 AA.
AC   P82322;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   UNKNOWN PROTEIN FROM 2D-PAGE OF THYLAKOID LUMEN (SPOT103)
DE   (FRAGMENT).
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Fabales; Fabaceae;
OC   Papilionoideae; Pisum.
RN   [1]
RP   SEQUENCE, SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.
RC   STRAIN=CV. DE GRACE; TISSUE=LEAF;
RA   Peltier J.B., Friso G., Kalume D.E., Roepstorff P., Nilsson F.,
RA   Adamska I., van Wijk K.J.;
RT   "Proteomics of the chloroplast.";
RL   Plant Cell 12:0-0(2000).
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE LUMEN.
CC   -!- DEVELOPMENTAL STAGE: UNFOLDED AND FULLY DEVELOPED LEAVES.
KW   Chloroplast; Thylakoid membrane.
FT   NON_TER      14      14
SQ   SEQUENCE   14 AA;  1381 MW;  0023DD7E0B97066B CRC64;


  Query Match             39.2%;  Score 20;  DB 10;  Length 14;
  Best Local Similarity   75.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 GSPY 9
          | ||
Db      4 GGPY 7


RESULT  16
Q9R599
ID   Q9R599      PRELIMINARY;      PRT;    15 AA.
AC   Q9R599;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   DNA TOPOISOMERASE I (FRAGMENT).
OS   Micrococcus luteus (Micrococcus lysodeikticus).
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Micrococcaceae; Micrococcus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93249439.
RA   Anderluzzi D., Pedrini A.M.;
RT   "Structural similarities between M. luteus and E. coli DNA
RT   topoisomerase I.";
RL   Biochem. Biophys. Res. Commun. 192:657-664(1993).
SQ   SEQUENCE   15 AA;  1680 MW;  C6A340B570A4CEE6 CRC64;


  Query Match             39.2%;  Score 20;  DB 2;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 2e+03;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 GVGSPYV 10
          |   |||
Db      2 GXYGPYV 8


RESULT  17
Q9UC85
ID   Q9UC85      PRELIMINARY;      PRT;    15 AA.
AC   Q9UC85;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   ERYTHROCYTE-DERIVED GROWTH-PROMOTING FACTOR (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95188211.
RA   Takeuchi A., Miyamoto T., Yamaji K., Masuho Y., Hayashi M.,
RA   Hayashi H., Onozaki K.;
RT   "A human erythrocyte-derived growth-promoting factor with a wide
RT   target cell spectrum: identification as catalase.";
RL   Cancer Res. 55:1586-1589(1995).
SQ   SEQUENCE   15 AA;  1415 MW;  C1E1EC38E9BE4C78 CRC64;


  Query Match             39.2%;  Score 20;  DB 4;  Length 15;
  Best Local Similarity   60.0%;  Pred. No. 2e+03;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 GVGSP 8
          | |:|
Db      7 GAGNP 11


RESULT  18
Q9S8F1
ID   Q9S8F1      PRELIMINARY;      PRT;    15 AA.
AC   Q9S8F1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   GLUTATHIONE S-TRANSFERASE ISOFORM II (EC 2.5.1.18) (FRAGMENT).
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95322859.
RA   Holt D.C., Lay V.J., Clarke E.D., Dinsmore A., Jepson I., Bright S.W.,
RA   Greenland A.J.;
RT   "Characterization of the safener-induced glutathione S-transferase
RT   isoform II from maize.";
RL   Planta 196:295-302(1995).
SQ   SEQUENCE   15 AA;  1530 MW;  2F105C48F7DD3A56 CRC64;


  Query Match             39.2%;  Score 20;  DB 10;  Length 15;
  Best Local Similarity   60.0%;  Pred. No. 2e+03;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 GVGSP 8
          | |:|
Db      8 GAGAP 12


RESULT  19
Q86869
ID   Q86869      PRELIMINARY;      PRT;    15 AA.
AC   Q86869;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   S-RNA PRODUCT, S-RNA PRODUCT (FRAGMENT).
```

```
OS   Lymphocytic choriomeningitis virus.
OC   Viruses; ssRNA negative-strand viruses; Arenaviridae; Arenavirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95190990.
RA   Moskophidis D., Zinkernagel R.M.;
RT   "Immunobiology of cytotoxic T-cell escape mutants of lymphocytic
RT   choriomeningitis virus.";
RL   J. Virol. 69:2187-2193(1995).
DR   EMBL; S75753; AAB33673.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   15 AA;  1571 MW;  2D25ABF4F776C1A7 CRC64;


  Query Match             39.2%;  Score 20;  DB 12;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 AGVGSP 8
           :|| ||
Db       2 SGVESP 7


RESULT  20
Q9TT30
ID   Q9TT30      PRELIMINARY;      PRT;    18 AA.
AC   Q9TT30;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   BETA-GLOBIN (FRAGMENT).
OS   Saguinus labiatus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BLOOD;
RA   Francino M.P., Ochman H.;
RT   "Strand Symmetry around the Beta-Globin Origin of Replication in
RT   Primates.";
RL   Mol. Biol. Evol. 0:0-0(2000).
DR   EMBL; AF205414; AAF23765.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   18 AA;  1940 MW;  13D07D8AAA1BFB37 CRC64;


  Query Match             39.2%;  Score 20;  DB 6;  Length 18;
  Best Local Similarity   57.1%;  Pred. No. 2.5e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 VMAGVGS 7
           |:||| :
Db       5 VVAGVAN 11


RESULT  21
Q9T2J5
ID   Q9T2J5      PRELIMINARY;      PRT;    18 AA.
AC   Q9T2J5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   5.6 KDA PHOTOSYSTEM I PSAK PROTEIN (FRAGMENT).
OS   Nicotiana tabacum (Common tobacco).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC   Solanaceae; Nicotiana.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 94105345.
RA   Obokata J., Mikami K., Hayashida N., Nakamura M., Sugiura M.;
RT   "Molecular heterogeneity of photosystem I. psaD, psaE, psaF, psaH, and
RT   psaL are all present in isoforms in Nicotiana spp.";
RL   Plant Physiol. 102:1259-1267(1993).
SQ   SEQUENCE   18 AA;  1950 MW;  EB628B87F4F73A07 CRC64;


  Query Match             39.2%;  Score 20;  DB 8;  Length 18;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 VGSP 8
           :|||
Db       3 IGSP 6


RESULT  22
Q62532
ID   Q62532      PRELIMINARY;      PRT;    18 AA.
AC   Q62532;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VA (EC 1.9.3.1) (FRAGMENT).
GN   COX5A.
OS   Mus spretus (Western wild mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRET/EI;
RX   MEDLINE; 94319082.
RA   Ko M.S., Wang X., Horton J.H., Hagen M.D., Takahashi N., Maezaki Y.,
RA   Nadeau J.H.;
RT   "Genetic mapping of 40 cDNA clones on the mouse genome by PCR.";
RL   Mamm. Genome 5:349-355(1994).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- FUNCTION: THIS IS THE HEME A-CONTAINING CHAIN.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O + 4
CC       FERRICYTOCHROME C.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
DR   EMBL; U05699; AAB60466.1; -.
DR   MGD; MGI:88474; Cox5a.
KW   Oxidoreductase; Heme; Mitochondrion.
FT   NON_TER       1      1
FT   VARIANT       9      9       S -> T.
SQ   SEQUENCE   18 AA;  1914 MW;  9533E5E252B20256 CRC64;


  Query Match             39.2%;  Score 20;  DB 11;  Length 18;
  Best Local Similarity   60.0%;  Pred. No. 2.5e+03;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 GVGSP 8
           |: ||
Db       6 GISSP 10


RESULT  23
Q91380
ID   Q91380      PRELIMINARY;      PRT;    18 AA.
AC   Q91380;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   RETINOID-X-RECEPTOR-GAMMA (FRAGMENT).
GN   RXR-<GAMMA>.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE; 94311845.
RA   Seleiro E.A., Darling D., Brickell P.M.;
RT   "The chicken retinoid-X-receptor-gamma gene gives rise to two distinct
RT   species of mRNA with different patterns of expression.";
RL   Biochem. J. 301:283-288(1994).
DR   EMBL; S72435; AAB31348.1; -.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2008 MW;  3AA890A5F97CF5C9 CRC64;


  Query Match             39.2%;  Score 20;  DB 13;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 GVGSPY 9
          |: :||
Db      4 GMQAPY 9


RESULT  24
Q9S8E2
ID   Q9S8E2      PRELIMINARY;      PRT;    19 AA.
AC   Q9S8E2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   6.1 KDA NUCLEAR-ENCODED PHOTOSYSTEM II REACTION CENTER SUBUNIT
DE   (FRAGMENT).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Caryophyllidae; Caryophyllales;
OC   Chenopodiaceae; Spinacia.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95340559.
RA   Irrgang K.D., Shi L.X., Funk C., Schroder W.P.;
RT   "A nuclear-encoded subunit of the photosystem II reaction center.";
RL   J. Biol. Chem. 270:17588-17593(1995).
SQ   SEQUENCE   19 AA;  2067 MW;  547B56337B5719E7 CRC64;


  Query Match             39.2%;  Score 20;  DB 10;  Length 19;
  Best Local Similarity   50.0%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GVGSPY 9
          | | |:
Db     10 GTGLPF 15


RESULT  25
Q9QUY4
ID   Q9QUY4      PRELIMINARY;      PRT;    19 AA.
AC   Q9QUY4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   OLIGODENDROCYTE-SPECIFIC UDP-GALACTOSE:CERAMIDE
DE   GALACTOSYLTRANSFERASE (FRAGMENT).
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96085162.
RA   Schulte S., Stoffel W.;
RT   "UDP galactose:ceramide galactosyltransferase and glutamate/aspartate
RT   transporter. Copurification, separation and characterization of the
RT   two glycoproteins.";
RL   Eur. J. Biochem. 233:947-953(1995).
SQ   SEQUENCE   19 AA;  1995 MW;  0FD4BAE303B99454 CRC64;


  Query Match             39.2%;  Score 20;  DB 11;  Length 19;
  Best Local Similarity   66.7%;  Pred. No. 2.6e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 AGVGSP 8
          | ||:|
Db     12 AQVGAP 17


RESULT  26
Q9PS70
ID   Q9PS70      PRELIMINARY;      PRT;    19 AA.
AC   Q9PS70;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   LOW DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92011685.
RA   Stifani S., Barber D.L., Aebersold R., Steyrer E., Shen X., Nimpf J.,
RA   Schneider W.J.;
RT   "The laying hen expresses two different low density lipoprotein
RT   receptor-related proteins.";
RL   J. Biol. Chem. 266:19079-19087(1991).
SQ   SEQUENCE   19 AA;  1861 MW;  4EEC931205620608 CRC64;


  Query Match             39.2%;  Score 20;  DB 13;  Length 19;
  Best Local Similarity   60.0%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 GVGSP 8
          |:| |
Db      5 GLGXP 9


RESULT  27
Q9UQW0
ID   Q9UQW0      PRELIMINARY;      PRT;     9 AA.
AC   Q9UQW0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PROLACTIN PRECURSOR (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 84182507.
RA   Truong A.T., Duez C., Belayew A., Renard A., Pictet R., Bell G.I.,
RA   Martial J.A.;
RT   "Isolation and characterization of the human prolactin gene.";
RL   EMBO J. 3:429-437(1984).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93076813.
RA   Peers B., Nalda A.M., Monget P., Voz M.L., Belayew A., Martial J.A.;
RT   "Binding of a 100-kDa ubiquitous factor to the human prolactin
RT   promoter is required for its basal and hormone-regulated activity.";
RL   Eur. J. Biochem. 210:53-58(1992).
DR   EMBL; X00368; CAA25108.1; -.
KW   Signal.
FT   SIGNAL        1      8       POTENTIAL.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1060 MW;  0A1A6775B8733054 CRC64;
```

```
  Query Match             37.3%;  Score 19;  DB 4;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 GSPY 9
          |||:
Db      5 GSPW 8


RESULT  28
Q9RFZ4
ID   Q9RFZ4      PRELIMINARY;      PRT;    13 AA.
AC   Q9RFZ4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   HYPOTHETICAL 1.5 KDA PROTEIN (FRAGMENT).
OS   Mycoplasma mycoides capri.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   capricolum group.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PG3;
RA   Thiaucourt F., Lorenzon S., David A.;
RT   "Phylogeny of the Mycoplasma mycoides cluster as shown by sequencing
RT   of a putative membrane protein gene.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF162998; AAF15253.1; -.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   13 AA;  1505 MW;  0B79431F5635573B CRC64;


  Query Match             37.3%;  Score 19;  DB 2;  Length 13;
  Best Local Similarity   60.0%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 GSPYV 10
          | ||:
Db      3 GIPYL 7


RESULT  29
Q9TRG9
ID   Q9TRG9      PRELIMINARY;      PRT;    15 AA.
AC   Q9TRG9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   ALKALINE PHOSPHODIESTERASE I (EC 3.1.4.1) (FRAGMENT).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RA   Maruyama E., Iwamatsu A., Takashima S.;
RL   Biochem. Mol. Biol. Int. 29:579-586(1993).
SQ   SEQUENCE   15 AA;  1720 MW;  C4FF771EBDC867E1 CRC64;


  Query Match             37.3%;  Score 19;  DB 6;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 3.1e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GVGSPYV 10
          || | |:
Db      4 GVHSQYL 10


RESULT  30
Q9TNQ3
```

```
ID   Q9TNQ3      PRELIMINARY;      PRT;    15 AA.
AC   Q9TNQ3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CLASS II HLA DR5 LIGAND.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 94164692.
RA   Falk K., Rotzschke O., Stevanovic S., Jung G., Rammensee H.G.;
RT   "Pool sequencing of natural HLA-DR, DQ, and DP ligands reveals
RT   detailed peptide motifs, constraints of processing, and general
RT   rules.";
RL   Immunogenetics 39:230-242(1994).
KW   MHC.
SQ   SEQUENCE   15 AA;  1738 MW;  5C8F3CE934481042 CRC64;


  Query Match             37.3%;  Score 19;  DB 7;  Length 15;
  Best Local Similarity   75.0%;  Pred. No. 3.1e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 SPYV 10
          :|||
Db     10 NPYV 13


RESULT  31
Q9UC53
ID   Q9UC53      PRELIMINARY;      PRT;    16 AA.
AC   Q9UC53;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   77 KDA SPONTANEOUS RECURRENT ABORTION-ASSOCIATED HUMAN EMBRYONIC
DE   ANTIGEN/IGVHIII HOMOLOG (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96033130.
RA   Shiraishi Y., Shiraishi Y., Yamamoto D., Hasegawa T., Kitamura W.,
RA   Miki S., Tanaka T., Suzuki T., Soma H.;
RT   "Diagnostic relevance of abortion-associated human embryonic antigen
RT   expressed on the cell surface of tumour promoter-treated Bloom
RT   syndrome cells.";
RL   Hum. Reprod. 10:1694-1701(1995).
SQ   SEQUENCE   16 AA;  1626 MW;  C9C5ED2512FF3FB9 CRC64;


  Query Match             37.3%;  Score 19;  DB 4;  Length 16;
  Best Local Similarity   66.7%;  Pred. No. 3.3e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 VMAGVG 6
          | :|||
Db      5 VESGVG 10


RESULT  32
Q9T2Q4
ID   Q9T2Q4      PRELIMINARY;      PRT;    16 AA.
AC   Q9T2Q4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CHAPERONIN-60 LS3 FRAGMENT.
OS   Brassica napus (Rape).
OG   Mitochondrion.
```

```
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales;
OC   Brassicaceae; Brassica.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 94302168.
RA   Cloney L.P., Bekkaoui D.R., Feist G.L., Lane W.S., Hemmingsen S.M.;
RT   "Brassica napus plastid and mitochondrial chaperonin-60 proteins
RT   contain multiple distinct polypeptides.";
RL   Plant Physiol. 105:233-241(1994).
SQ   SEQUENCE   16 AA;  1901 MW;  CFAE799B7C938063 CRC64;


  Query Match             37.3%;  Score 19;  DB 8;  Length 16;
  Best Local Similarity   66.7%;  Pred. No. 3.3e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GVGSPY 9
           |  |||
Db       5 GYISPY 10


RESULT  33
Q9T2V8
ID   Q9T2V8      PRELIMINARY;      PRT;    16 AA.
AC   Q9T2V8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   2-ENOYL-COA HYDRATASE (FRAGMENT).
OS   Homo sapiens (Human).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95046784.
RA   Middleton B.;
RT   "The mitochondrial long-chain trifunctional enzyme: 2-enoyl-CoA
RT   hydratase, 3-hydroxyacyl-CoA dehydrogenase and 3-oxoacyl-CoA
RT   thiolase.";
RL   Biochem. Soc. Trans. 22:427-431(1994).
SQ   SEQUENCE   16 AA;  1763 MW;  31AD66A3080B019A CRC64;


  Query Match             37.3%;  Score 19;  DB 8;  Length 16;
  Best Local Similarity   50.0%;  Pred. No. 3.3e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 VMAGVGSP 8
           |: || :|
Db       9 VVDGVRTP 16


RESULT  34
Q07055
ID   Q07055      PRELIMINARY;      PRT;    17 AA.
AC   Q07055;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   P18 PROTEIN (FRAGMENT).
OS   Crithidia fasciculata.
OG   Mitochondrion.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Crithidia.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93189582.
RA   Xu C., Ray D.S.;
RT   "Isolation of proteins associated with kinetoplast DNA networks in
RT   vivo.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:1786-1789(1993).
DR   EMBL; S56494; AAB25703.1; -.
KW   Kinetoplast.
FT   NON_TER      17      17
SQ   SEQUENCE   17 AA;  2064 MW;  FEB4A61689366B59 CRC64;


  Query Match             37.3%;  Score 19;  DB 8;  Length 17;
  Best Local Similarity   75.0%;  Pred. No. 3.5e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 SPYV 10
           |||:
Db      12 SPYM 15


RESULT  35
Q9S882
ID   Q9S882      PRELIMINARY;      PRT;    19 AA.
AC   Q9S882;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   ATP SYNTHASE SUBUNIT II-B' (FRAGMENT).
OS   Chlamydomonas reinhardtii.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC   Chlamydomonadaceae; Chlamydomonas.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96128220.
RA   Fiedler H.R., Schmid R., Leu S., Shavit N., Strotmann H.;
RT   "Isolation of CF0CF1 from Chlamydomonas reinhardtii cw15 and the N-
RT   terminal amino acid sequences of the CF0CF1 subunits.";
RL   FEBS Lett. 377:163-166(1995).
SQ   SEQUENCE   19 AA;  2081 MW;  A0AC64A247D406A2 CRC64;


  Query Match             37.3%;  Score 19;  DB 10;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 4e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VMAG 4
           ||||
Db      13 VMAG 16


Search completed: December 12, 2000, 02:44:51
Job time: 5023 sec
```

```
OM of: US-08-860-232-12 to: EST:*   out_format : pfs

Date: Dec 12, 2000  3:01 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US08860232/runat_11122000_153023_14112/app_query.fasta_1.68
 -DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US08860232_@CGN1_1_1781 -NCPU=6
 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-08-860-232-12
Query length: 10
Database: EST:*
Database sequences: 7189864
Database length: -1203564053
Search time (sec): 931.330000



score_list:
Sequence              Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_est33:BE091535      -    51.00  165.30  3.05       590  ! BE091535 PM0-BT0730-280300-001-
gb_est22:AW414405      +    51.00  164.85  3.24       624  ! AW414405 uo98h12.y1 NCI_CGAP_Ma
gb_est13:AI905418      +    51.00  163.90  3.65       701  ! AI905418 RC-BT091-200199-085 BT
gb_gss9:AQ625244       +    41.00  131.56  231.36     531  ! AQ625244 CITBI-E1-2650N19.TF CI
gb_gss8:AQ575993       +    41.00  129.92  285.52     650  ! AQ575993 nbxb0088K20f CUGI Rice
gb_est36:D29356        +    40.00  134.14  166.06     252  ! D29356 HUMNK454 Human epidermal
gb_est15:AV032124      -    39.00  133.10  189.76     187  ! AV032124 AV032124 Mus musculus
gb_est8:AI097982       -    39.00  132.85  196.10     193  ! AI097982 vg82c05.r1 Barstead MF
gb_est5:AA645141       -    39.00  130.84  253.47     247  ! AA645141 vs72f02.r1 Stratagene
gb_est11:AI626890      +    39.00  130.81  254.54     248  ! AI626890 vs72f02.x1 Stratagene
gb_est11:AI592362      -    39.00  130.55  263.09     256  ! AI592362 vs72f02.y1 Stratagene
gb_est28:BB246984      +    39.00  129.49  301.68     292  ! BB246984 RIKEN full-le
gb_est12:AI648664      +    39.00  129.32  308.13     298  ! AI648664 tx63g11.x1 NCI_CGAP_Ut
gb_est20:AW197689      -    39.00  129.19  313.51     303  ! AW197689 xm85h01.x1 NCI_CGAP_Ki
gb_est18:AV413366      -    39.00  127.68  380.51     365  ! AV413366 AV413366 Lotus japonic
gb_gss11:AQ811654      -    39.00  127.66  381.59     366  ! AQ811654 HS_5460_B1_H03_SP6E RP
gb_est18:AV423884      -    39.00  127.54  387.02     371  ! AV423884 AV423884 Lotus japonic
gb_est8:AI117420       -    39.00  127.04  413.09     395  ! AI117420 ub78g08.r1 Soares_mamm
gb_gss11:AQ772784      +    39.00  126.71  430.52     411  ! AQ772784 HS_2019_B1_H10_T7C CIT
gb_est18:AV427332      -    39.00  126.58  438.15     418  ! AV427332 AV427332 Lotus japonic
gb_est11:AI614514      -    39.00  126.52  441.42     421  ! AI614514 vh59h06.y1 Soares_mamm
gb_est25:AW986801      -    39.00  126.18  461.07     439  ! AW986801 uf87g12.y1 Soares_mamm
gb_gss5:AQ388517       +    39.00  125.59  497.18     472  ! AQ388517 RPCI11-153H3.TV RPCI-1
gb_gss14:AQ982686      -    39.00  125.52  501.56     476  ! AQ982686 RPCI-23-307D10.TV RPCI
gb_est4:AA510558       -    39.00  124.82  548.78     519  ! AA510558 vh59h06.r1 Soares_mamm
gb_gss10:AQ728333      +    39.00  124.79  550.98     521  ! AQ728333 HS_5470_A1_B12_T7A RPC
gb_est6:AA796380       -    39.00  124.67  559.79     529  ! AA796380 vs99e03.r1 Barstead mc
gb_gss14:AQ982968      -    39.00  123.99  610.51     575  ! AQ982968 RPCI-23-307F10.TV RPCI
gb_est21:AW349898      -    39.00  121.46  845.13     786  ! AW349898 GM210006A20F3R Gm-r102
gb_est34:BE283122      -    39.00  121.16  877.60     815  ! BE283122 601103581F1 NCI_CGAP_L
gb_est33:BE063147      -    38.00  123.58  643.83     395  ! BE063147 CM2-BT0266-221199-037-
gb_est37:F15107        +    38.00  123.23  672.69     412  ! F15107 SSC9F02 Porcine small in
gb_est5:AA690407       -    38.00  122.70  720.31     440  ! AA690407 vu52c01.r1 Soares_mamm
gb_est5:AA583896       +    38.00  122.57  732.24     447  ! AA583896 nn64a10.s1 NCI_CGAP_La
gb_gss5:AQ334126       -    38.00  121.46  845.03     513  ! AQ334126 HS_5010_A1_H11_T7 RPCI
gb_gss2:AQ164697       -    38.00  121.44  846.74     514  ! AQ164697 HS_3006_B2_D08_T7 CIT
gb_est14:AL367987      -    38.00  121.36  855.32     519  ! AL367987 MtBA21E08F1 MtBA Medic
gb_gss14:AQ996450      -    38.00  121.36  855.32     519  ! AQ996450 RPCI-23-387B20.TV RPCI
gb_gss8:AQ577192       +    38.00  120.97  899.94     545  ! AQ577192 nbxb0090H09f CUGI Rice
gb_est34:BE203415      -    38.00  119.96  1.0e+03    617  ! BE203415 EST403437 KV1 Medicago
gb_gss5:AQ364822       +    38.00  119.44  1.1e+03    658  ! AQ364822 nbxb0061L23r CUGI Rice
gb_est23:AW686204      -    38.00  119.36  1.1e+03    664  ! AW686204 NF035B10NR1F1000 No
gb_est34:BE248924      -    38.00  119.19  1.1e+03    678  ! BE248924 NF023H12DT1F1101 Dr
gb_gss21:AG013496      -    38.00  118.66  1.2e+03    724  ! AG013496 Homo sapiens genomi
gb_gss17:AZ196516      +    38.00  118.61  1.2e+03    729  ! AZ196516 SP_1032_A1_E06_SP6E

seq_name: gb_est33:BE091535

seq_documentation_block:
LOCUS       BE091535      590 bp    mRNA            EST       12-JUN-2000
DEFINITION  PM0-BT0730-280300-001-b10 BT0730 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE091535
VERSION     BE091535.1  GI:8481987
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=PM0-BT0730-280
            300-001-b10&t3=2000-03-28&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 6
            High quality sequence stop: 577.
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BT0730"
                     /dev_stage="Adult"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT      142 a    161 c    151 g    135 t      1 others
ORIGIN


alignment_scores:
           Quality:   51.00            Length:     10
             Ratio:  5.100              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-12 x BE091535/rev   ..

 Align seg 1/1  to reverse of: BE091535  from: 1  to: 590

       1 ValMetAlaGlyValGlySerProTyrVal 10
         ||||||||||||||||||||||||||||||
     505 GTGATGGCTGGTGTGGGCTCCCCATATGTC 476
```

```
seq_name: gb_est22:AW414405

seq_documentation_block:
LOCUS       AW414405      624 bp    mRNA            EST       09-FEB-2000
DEFINITION  uo98h12.y1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:2650631 5'
            similar to gb:M11730 ERBB-2 RECEPTOR PROTEIN-TYROSINE KINASE
            PRECURSOR (HUMAN); gb:X78987 M.musculus (MOUSE);, mRNA sequence.
ACCESSION   AW414405
VERSION     AW414405.1  GI:6940731
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            MGL:1031083
            Seq primer: -40RP from Gibco
            High quality sequence stop: 427.
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="Mus musculus"
                     /strain="129 - C57/B6 - FVBN"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:2650631"
                     /clone_lib="NCI_CGAP_Mam3"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="10 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
                     Reference for transgenic model: Xu et al., Nature Genetics
                     22, 37-43 (1999)."
BASE COUNT      140 a    161 c    179 g    144 t
ORIGIN


alignment_scores:
           Quality:   51.00            Length:     10
             Ratio:  5.100              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-12 x AW414405   ..

 Align seg 1/1  to: AW414405  from: 1  to: 624

       1 ValMetAlaGlyValGlySerProTyrVal 10
         ||||||||||||||||||||||||||||||
      27 GTCATGGCTGGTGTGGGTTCTCCATATGTG 56

seq_name: gb_est13:AI905418

seq_documentation_block:
LOCUS       AI905418      701 bp    mRNA            EST       30-MAR-2000
DEFINITION  RC-BT091-200199-085 BT091 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI905418
VERSION     AI905418.1  GI:6495805
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 701)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=RC&t2=RC-BT091-085.html
            &t3=200199&t4=1)
            Seq primer: puc 18 forward.
FEATURES             Location/Qualifiers
     source          1. .701
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BT091"
                     /sex="female"
                     /dev_stage="Adult"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT      155 a    181 c    208 g    148 t      9 others
ORIGIN


alignment_scores:
           Quality:   51.00            Length:     10
             Ratio:  5.100              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-12 x AI905418   ..

 Align seg 1/1  to: AI905418  from: 1  to: 701

       1 ValMetAlaGlyValGlySerProTyrVal 10
         ||||||||||||||||||||||||||||||
     373 GTGATGGCTGGTGTGGGCTCCCCATATGTC 402

seq_name: gb_gss9:AQ625244

seq_documentation_block:
LOCUS       AQ625244      531 bp    DNA             GSS       16-JUN-1999
DEFINITION  CITBI-E1-2650N19.TF CITBI-E1 Homo sapiens genomic clone 2650N19,
            DNA sequence.
ACCESSION   AQ625244
VERSION     AQ625244.1  GI:5087636
KEYWORDS    GSS.
SOURCE      human.
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 531)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CITBI-E1-2650N19.TR
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .531
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2650N19"
                     /clone_lib="CITBI-E1"
                     /sex="male"
                     /cell_type="sperm"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"
BASE COUNT      127 a    126 c    107 g    169 t      2 others
ORIGIN


alignment_scores:
            Quality:   41.00             Length:      8
              Ratio:  5.125               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 87.500

alignment_block:
 US-08-860-232-12 x AQ625244   ..

 Align seg 1/1  to: AQ625244  from: 1  to: 531

       3 AlaGlyValGlySerProTyrVal 10
         ||||||:::|||||||||||||||
     503 GCAGGAATTGGTAGCCCTTATGTG 526

seq_name: gb_gss8:AQ575993

seq_documentation_block:
LOCUS       AQ575993      650 bp    DNA             GSS       02-JUN-1999
DEFINITION  nbxb0088K20f CUGI Rice BAC Library Oryza sativa genomic clone
            nbxb0088K20f, DNA sequence.
ACCESSION   AQ575993
VERSION     AQ575993.1  GI:4976478
KEYWORDS    GSS.
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
REFERENCE   1  (bases 1 to 650)
  AUTHORS   Wing,R.A. and Dean,R.A.
  TITLE     A BAC End Sequencing Framework to Sequence the Rice Genome
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            Class: BAC ends
            High quality sequence stop: 454.
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Oryza sativa"
                     /strain="Japonica"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:4530"
                     /clone="nbxb0088K20f"
                     /clone_lib="CUGI Rice BAC Library"
                     /tissue_type="Leaf"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII; Rice is one of two most popular grains in the
                     world. Half of the world  population especially those
                     inhabiting highly populated areas of the  humid tropics
                     and subtropics, rely on rice as their primary source of
                     carbohydrate. Monocotyledonous rice is a diploid plant
                     (2n=24) with a  haploid genome equivalent of 431 Mbp
                     (Arumuganathan and Earle, 1991).  The relatively small
                     genome of rice, three times larger than that of
                     Arabidopsis, makes it suitable for genomic studies. In
                     order to  facilitate positional cloning, physical mapping
                     and genome sequencing  of rice, we have constructed a BAC
                     library from Oryza sativa, Nipponbare  variety. The
                     library contains 36,864 clones with an average insert size
                     of 128.5 Kb providing 10.9 haploid genome equivalents. The
                     deep  coverage allows the isolation a particular sequence
                     with a probability  of 99.9 %. Two high density filters,
                     each containing 18,432 clones  (doubly spotted), represent
                     the whole library for colony screening."
BASE COUNT      226 a    100 c    129 g    194 t      1 others
ORIGIN


alignment_scores:
            Quality:   41.00             Length:     10
              Ratio:  4.556               Gaps:      0
 Percent Similarity:  90.000   Percent Identity: 70.000

alignment_block:
 US-08-860-232-12 x AQ575993   ..

 Align seg 1/1  to: AQ575993  from: 1  to: 650

       1 ValMetAlaGlyValGlySerProTyrVal 10
         :::|||||||||   ||||||||||||:::
     475 ATGATGGCAGGTCAGGGCTCTCCCTACATA 504

seq_name: gb_est36:D29356

seq_documentation_block:
LOCUS       D29356        252 bp    mRNA            EST       14-NOV-1997
DEFINITION  HUMNK454 Human epidermal keratinocyte Homo sapiens cDNA clone 454,
            mRNA sequence.
ACCESSION   D29356
VERSION     D29356.1  GI:599289
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 252)
  AUTHORS   Konishi,K., Morishima,Y.-I., Ueda,E., Nonomura,K., Kibe,S.,
            Yamanishi,K. and Yasuno,H.
  TITLE     Cataloging of the genes expressed in human keratinocytes: analysis
            of 607 randomly isolated cDNA sequences
  JOURNAL   Biochem. Biophys. Res. Commun. 202, 976-983 (1994)
  MEDLINE   94324994
COMMENT     Contact: Kiyofumi Yamanishi
            Department of Dermatology
```

```
            Kyoto Prefectural University of Medicine
            Kamigyo-ku, Kyoto, Kyoto, 602, Japan
            Tel: 75-251-5587
            Fax: 75-251-5587.
FEATURES              Location/Qualifiers
     source           1. .252
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="454"
                      /clone_lib="Human epidermal keratinocyte"
                      /note="tissue_type=Epidermis, cell_type=Keratinocyte,
                      cell_line=primary culture"
BASE COUNT       50 a     69 c     56 g     71 t      6 others
ORIGIN


alignment_scores:
            Quality:   40.00             Length:     10
              Ratio:  4.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-12 x D29356   ..

 Align seg 1/1  to: D29356  from: 1  to: 252

       1 ValMetAlaGlyValGlySerProTyrVal 10
         ::::::::::|||:::|||||||||||||||
     115 ATCTTTGGGGGAATCGGTTCCCCATATGTC 144

seq_name: gb_est15:AV032124

seq_documentation_block:
LOCUS        AV032124      187 bp    mRNA            EST       31-AUG-1999
DEFINITION   AV032124 Mus musculus adult C57BL/6J cerebellum Mus musculus cDNA
             clone 1500034B14, mRNA sequence.
ACCESSION    AV032124
VERSION      AV032124.1  GI:4831672
KEYWORDS     EST.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 187)
  AUTHORS    Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K.,
             Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T., Hara
             ,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M., Kawai,J.,
             Kikuchi,N., Kojima,Y., Matsuyama,T., Niitsuma,H., Oda,H., Owa,C.,
             Sato,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara
             ,Y., Suzuki,H., Suzuki,H., Tateno,M., Tomaru,Y., Tominaga,N.,
             Watanabe,S., Yagame,M., Yamamura,T., Yokota,T., Yoshino,M.,
             Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
  TITLE      RIKEN Mouse ESTs
  JOURNAL    Unpublished (1999)
COMMENT      Contact: Chie Owa
             Genome Science Laboratory
             RIKEN
             3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
             Tel: 81-298-36-9145
             Fax: 81-298-36-9098
             Email: genome-res@rtc.riken.go.jp
             Thermostabilization and thermoactivation of thermolabile enzymes by
             trehalose and its application for the synthesis of full length cDNA
             (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
             Transcriptional sequencing: A method for DNA sequencing using RNA
             polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
             Please visit our web site (http://genome.rtc.riken.go.jp) for
             further details.
FEATURES              Location/Qualifiers
     source           1. .187
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="1500034B14"
                      /clone_lib="Mus musculus adult C57BL/6J cerebellum"
                      /sex="male"
                      /tissue_type="cerebellum"
                      /dev_stage="adult"
BASE COUNT       48 a     69 c     34 g     36 t
ORIGIN


alignment_scores:
            Quality:   39.00             Length:     10
              Ratio:  4.333               Gaps:      0
 Percent Similarity:  90.000   Percent Identity: 80.000

alignment_block:
 US-08-860-232-12 x AV032124/rev   ..

 Align seg 1/1  to reverse of: AV032124  from: 1  to: 187

       1 ValMetAlaGlyValGlySerProTyrVal 10
         |||:::|||||||||||||||||||   |||
      40 GTCCTTGCTGGGGTGGGTAGTCCATGTGTA 11

seq_name: gb_est8:AI097982

seq_documentation_block:
LOCUS        AI097982      193 bp    mRNA            EST       20-AUG-1998
DEFINITION   vg82c05.r1 Barstead MPLRB1 Mus musculus cDNA clone IMAGE:872456 5',
             mRNA sequence.
ACCESSION    AI097982
VERSION      AI097982.1  GI:3447507
KEYWORDS     EST.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 193)
  AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
             Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
             Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
             Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
             Waterston,R.
  TITLE      The WashU-HHMI Mouse EST Project
  JOURNAL    Unpublished (1996)
COMMENT      Contact: Marra M/Mouse EST Project
             WashU-HHMI Mouse EST Project
             Washington University School of MedicineP
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:511936
             Seq primer: -28m13 rev2 ET from Amersham
             High quality sequence stop: 145.
FEATURES              Location/Qualifiers
     source           1. .193
                      /organism="Mus musculus"
                      /strain="BALB/c"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:872456"
                      /clone_lib="Barstead MPLRB1"
                      /sex="mixed"
                      /tissue_type="Kidney"
                      /dev_stage="6 weeks"
                      /lab_host="DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      [CATGGATTCGGTACC], digested with Not I and cloned into the
```

```
                    Not I and Eco RI sites of the modified pT7T3 vector.
                    Library constructed by Bob Barstead."
BASE COUNT       46 a      55 c      53 g      39 t
ORIGIN


alignment_scores:
            Quality:   39.00              Length:      10
              Ratio:  4.333                Gaps:       0
 Percent Similarity:  90.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-12 x AI097982/rev   ..

 Align seg 1/1  to reverse of: AI097982  from: 1  to: 193

       1 ValMetAlaGlyValGlySerProTyrVal 10
         |||   ||||||:::|||:::||||||:::
     137 GTTTCTGCTGGAATTGGAGGCCCATACATC 108

seq_name: gb_est5:AA645141

seq_documentation_block:
LOCUS        AA645141      247 bp    mRNA            EST       28-OCT-1997
DEFINITION   vs72f02.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
             IMAGE:1151835 5', mRNA sequence.
ACCESSION    AA645141
VERSION      AA645141.1  GI:2571570
KEYWORDS     EST.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 247)
  AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
             Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
             Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
             Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
             Waterston,R.
  TITLE      The WashU-HHMI Mouse EST Project
  JOURNAL    Unpublished (1996)
COMMENT      Contact: Marra M/Mouse EST Project
             WashU-HHMI Mouse EST Project
             Washington University School of MedicineP
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800 -
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:625043
             Seq primer: -28m13 rev1 ET from Amersham.
FEATURES              Location/Qualifiers
     source           1. .247
                      /organism="Mus musculus"
                      /strain="C57BL/6"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:1151835"
                      /clone_lib="Stratagene mouse skin (#937313)"
                      /sex="females"
                      /tissue_type="whole skin"
                      /dev_stage="11 weeks old"
                      /lab_host="SOLR (kanamycin resistant)"
                      /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                      ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                      dT. Whole skin from 11 week old C57BL/6 female mice.
                      Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                      adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                      sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       54 a      80 c      55 g      58 t
ORIGIN
```

```
alignment_scores:
            Quality:   39.00              Length:       9
              Ratio:  4.875                Gaps:       0
 Percent Similarity:  88.889   Percent Identity: 77.778

alignment_block:
 US-08-860-232-12 x AA645141/rev   ..

 Align seg 1/1  to reverse of: AA645141  from: 1  to: 247

       1 ValMetAlaGlyValGlySerProTyr 9
         |||:::|||||||||||||   ||||||
      38 GTTCTGGCAGGGGTGGGTCATCCATAC 12

seq_name: gb_est11:AI626890

seq_documentation_block:
LOCUS        AI626890      248 bp    mRNA            EST       23-APR-1999
DEFINITION   vs72f02.x1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
             IMAGE:1151835 3', mRNA sequence.
ACCESSION    AI626890
VERSION      AI626890.1  GI:4663690
KEYWORDS     EST.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 248)
  AUTHORS    Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
             Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
             ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
             ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
             Waterston,R. and Wilson,R.
  TITLE      The WashU-NCI Mouse EST Project 1999
  JOURNAL    Unpublished (1999)
COMMENT      Contact: Marra M/WashU-NCI Mouse EST Project 1999
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:625043
             This clone was previously sequenced on the 5' end only, this new
             data is from the 3' end.
FEATURES              Location/Qualifiers
     source           1. .248
                      /organism="Mus musculus"
                      /strain="C57BL/6"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:1151835"
                      /clone_lib="Stratagene mouse skin (#937313)"
                      /sex="females"
                      /tissue_type="whole skin"
                      /dev_stage="11 weeks old"
                      /lab_host="SOLR (kanamycin resistant)"
                      /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                      ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                      dT. Whole skin from 11 week old C57BL/6 female mice.
                      Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                      adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                      sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       59 a      54 c      81 g      54 t
ORIGIN


alignment_scores:
            Quality:   39.00              Length:       9
              Ratio:  4.875                Gaps:       0
 Percent Similarity:  88.889   Percent Identity: 77.778

alignment_block:
```

```
 US-08-860-232-12 x AI626890   ..

 Align seg 1/1  to: AI626890  from: 1  to: 248

       1 ValMetAlaGlyValGlySerProTyr 9
         |||:::|||||||||||||   ||||||
     220 GTTCTGGCAGGGGTGGGTCATCCATAC 246

seq_name: gb_est11:AI592362

seq_documentation_block:
LOCUS       AI592362      256 bp    mRNA            EST       21-APR-1999
DEFINITION  vs72f02.y1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
            IMAGE:1151835 5', mRNA sequence.
ACCESSION   AI592362
VERSION     AI592362.1  GI:4601410
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 256)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:625043
            This read is a RESEQUENCE of a previously sequenced mouse clone
            This read has been verified (found to hit its original self in the
            correct orientation)
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .256
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1151835"
                     /clone_lib="Stratagene mouse skin (#937313)"
                     /sex="females"
                     /tissue_type="whole skin"
                     /dev_stage="11 weeks old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. Whole skin from 11 week old C57BL/6 female mice.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       58 a     81 c     58 g     59 t
ORIGIN


alignment_scores:
           Quality:   39.00            Length:      9
             Ratio:  4.875              Gaps:      0
 Percent Similarity:  88.889   Percent Identity: 77.778

alignment_block:
 US-08-860-232-12 x AI592362/rev   ..

 Align seg 1/1  to reverse of: AI592362  from: 1  to: 256

       1 ValMetAlaGlyValGlySerProTyr 9
         |||:::|||||||||||||   ||||||
      48 GTTCTGGCAGGGGTGGGTCATCCATAC 22

seq_name: gb_est28:BB246984

seq_documentation_block:
LOCUS       BB246984      292 bp    mRNA            EST       06-JUL-2000
DEFINITION  BB246984 RIKEN full-length enriched, 7 days neonate cerebellum Mus
            musculus cDNA clone A730016O20 3', mRNA sequence.
ACCESSION   BB246984
VERSION     BB246984.1  GI:8939730
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 292)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
            ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.
            , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.
            , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.
            , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya
            ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.
            , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
            ,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Genome Exploration Research Group, Life Science Tsukuba Center,
            Genome Science Laboratory
            The Institute of Physical and Chemical Research (RIKEN), Genomic
            Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: +81-298-36-9013
            Fax: +81-298-36-9098
            Email: genome-res@rtc.riken.go.jp,
            URL:http://genome.rtc.riken.go.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .292
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A730016O20"
                     /clone_lib="RIKEN full-length enriched, 7 days neonate
                     cerebellum"
                     /tissue_type="cerebellum"
                     /dev_stage="7 days neonate"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
```

primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT 77 a 65 c 47 g 103 t
ORIGIN

alignment_scores:
Quality: 39.00 Length: 9
Ratio: 4.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 55.556

alignment_block:
US-08-860-232-12 x BB246984 ..

Align seg 1/1 to: BB246984 from: 1 to: 292

2 MetAlaGlyValGlySerProTyrVal 10
::::::|||:::|||||||||||||:::
76 CTAAGTGGCATTGGTTCGCCCTATATT 102

seq_name: gb_est12:AI648664

seq_documentation_block:
LOCUS AI648664 298 bp mRNA EST 16-DEC-1999
DEFINITION tx63g11.x1 NCI_CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2274308 3' similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION AI648664
VERSION AI648664.1 GI:4729498
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 298)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 860 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 277.
FEATURES Location/Qualifiers
source 1. .298
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2274308"
/clone_lib="NCI_CGAP_Ut1"
/tissue_type="well-differentiated endometrial adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Life Technologies catalog #: 11538-014"
BASE COUNT 70 a 71 c 61 g 96 t
ORIGIN

alignment_scores:
Quality: 39.00 Length: 10
Ratio: 4.333 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 60.000

alignment_block:
US-08-860-232-12 x AI648664 ..

Align seg 1/1 to: AI648664 from: 1 to: 298

1 ValMetAlaGlyValGlySerProTyrVal 10
:::::: |||:::|||||||||||||||
164 ATTTTTTTAGGGATTGGGTCTCCCTATGTT 193

seq_name: gb_est20:AW197689

seq_documentation_block:
LOCUS AW197689 303 bp mRNA EST 29-NOV-1999
DEFINITION xm85h01.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2691025 3' similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION AW197689
VERSION AW197689.1 GI:6476919
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 303)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco.
FEATURES Location/Qualifiers
source 1. .303
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2691025"
/clone_lib="NCI_CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 90 a 76 c 70 g 67 t
ORIGIN

alignment_scores:
Quality: 39.00 Length: 8
Ratio: 4.875 Gaps: 0

```
 Percent Similarity: 100.000   Percent Identity: 87.500

alignment_block:
 US-08-860-232-12 x AW197689/rev   ..

 Align seg 1/1  to reverse of: AW197689  from: 1  to: 303

       3 AlaGlyValGlySerProTyrVal 10
         :::||||||||||||||||||||
     253 AGTGGGGTGGGTTCTCCCTATGTT 230

seq_name: gb_est18:AV413366

seq_documentation_block:
LOCUS       AV413366      365 bp    mRNA            EST       23-MAY-2000
DEFINITION  AV413366 Lotus japonicus young plants (two-week old) Lotus
            japonicus cDNA clone MWM231d09_r 5', mRNA sequence.
ACCESSION   AV413366
VERSION     AV413366.1  GI:7742542
KEYWORDS    EST.
SOURCE      Lotus japonicus.
  ORGANISM  Lotus japonicus
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
            Fabales; Fabaceae; Papilionoideae; Lotus.
REFERENCE   1  (bases 1 to 365)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     Generation of 7137 non-redundant expressed sequence tags from a
            legume, Lotus japonicus
  JOURNAL   DNA Res. 7 (2), 127-130 (2000)
  MEDLINE   20277479
COMMENT     Contact: Yasukazu Nakamura
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: ynakamu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .365
                     /organism="Lotus japonicus"
                     /db_xref="taxon:34305"
                     /clone="MWM231d09_r"
                     /clone_lib="Lotus japonicus young plants (two-week old)"
                     /dev_stage="young plants (two-week old)"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI; isolate=Miyakojima MG-20"
BASE COUNT      108 a     74 c    100 g     83 t
ORIGIN


alignment_scores:
            Quality:   39.00             Length:      9
              Ratio:  4.333               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 66.667

alignment_block:
 US-08-860-232-12 x AV413366/rev   ..

 Align seg 1/1  to reverse of: AV413366  from: 1  to: 365

       2 MetAlaGlyValGlySerProTyrVal 10
         ::::::|||:::|||||||||||||||
     364 TTAACAGGAATAGGTTCACCGTACGTG 338
```

```
OM of: US-08-860-232-1 to: GenEmbl:*   out_format : pfs

Date: Dec 12, 2000  5:48 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US08860232/runat_11122000_154522_20735/app_query.fasta_1.67
 -DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US08860232_@CGN1_1_3727 -NCPU=6
 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-08-860-232-1
Query length: 11
Database: GenEmbl:*
Database sequences: 1033670
Database length: -2111177393
Search time (sec): 909.990000



score_list:
Sequence              Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_pr6:HSP53001         +   55.00  179.47  0.0713    735  ! X60010 Human mRNA for mutated p
gb_pat:A61761           +   55.00  176.41  0.1055   1098  ! A61761 Sequence 1 from Patent
gb_pr6:HSP53002         +   55.00  175.87  0.1131   1179  ! X60011 Human mRNA for mutated
gb_pr6:HSP53003         +   55.00  175.87  0.1131   1179  ! X60012 Human mRNA for mutated
gb_pr6:HSP53004         +   55.00  175.87  0.1131   1179  ! X60013 Human mRNA for mutated
gb_pr6:HSP53005         +   55.00  175.87  0.1131   1179  ! X60014 Human mRNA for mutated
gb_pr6:HSP53006         +   55.00  175.87  0.1131   1179  ! X60015 Human mRNA for mutated
gb_pr6:HSP53007         +   55.00  175.87  0.1131   1179  ! X60016 Human mRNA for mutated
gb_pr6:HSP53008         +   55.00  175.87  0.1131   1179  ! X60017 Human mRNA for mutated
gb_pr6:HSP53009         +   55.00  175.87  0.1131   1179  ! X60018 Human mRNA for mutated
gb_pr6:HSP53010         +   55.00  175.87  0.1131   1179  ! X60019 Human mRNA for mutated
gb_pr6:HSP53011         +   55.00  175.87  0.1131   1179  ! X60020 Human mRNA for mutated
gb_pat:A61360           +   55.00  175.85  0.1134   1182  ! A61360 Sequence 2 from Patent
gb_pat:AR052878         +   55.00  175.85  0.1134   1182  ! AR052878 Sequence 215 from pat
gb_pat:AR061772         +   55.00  175.85  0.1134   1182  ! AR061772 Sequence 92 from pate
gb_pat:AR061773         +   55.00  175.85  0.1134   1182  ! AR061773 Sequence 93 from pate
gb_pat:AR061774         +   55.00  175.85  0.1134   1182  ! AR061774 Sequence 94 from pate
gb_pat:AR062013         +   55.00  175.85  0.1134   1182  ! AR062013 Sequence 92 from pate
gb_pat:AR062014         +   55.00  175.85  0.1134   1182  ! AR062014 Sequence 93 from pate
gb_pat:AR062015         +   55.00  175.85  0.1134   1182  ! AR062015 Sequence 94 from pate
gb_pr7:U48956           +   55.00  175.85  0.1134   1182  ! U48956 Macaca mulatta p53 gene
gb_pat:I28717           +   55.00  175.64  0.1165   1215  ! I28717 Sequence 20 from patent
gb_pat:I89371           +   55.00  175.64  0.1165   1215  ! I89371 Sequence 20 from patent
gb_om:AF175893          +   55.00  175.28  0.1219   1273  ! AF175893 Tupaia belangeri chin
gb_pat:I22243           +   55.00  175.10  0.1247   1303  ! I22243 Sequence 13 from patent
gb_pr7:HUMTP53B         +   55.00  175.10  0.1247   1303  ! M14695 Human p53 cellular tumo
gb_pat:I22242           +   55.00  175.08  0.1251   1307  ! I22242 Sequence 12 from patent
gb_pr7:HUMTP53A         +   55.00  175.08  0.1251   1307  ! M14694 Human p53 cellular tumo
gb_pat:A67149           +   55.00  175.02  0.1260   1317  ! A67149 Sequence 6 from Patent
gb_pat:AR064393         +   55.00  175.02  0.1260   1317  ! AR064393 Sequence 1 from paten
gb_pat:I28707           +   55.00  175.02  0.1260   1317  ! I28707 Sequence 1 from patent
gb_pat:I89361           +   55.00  175.02  0.1260   1317  ! I89361 Sequence 1 from patent
gb_pr6:HSP53            +   55.00  175.02  0.1260   1317  ! X02469 Human mRNA for p53 cell
gb_pr4:CAP53            +   55.00  174.56  0.1338   1400  ! X16384 African Green Monkey mR
gb_pat:AR052767         +   55.00  174.11  0.1416   1484  ! AR052767 Sequence 103 from pat
gb_pat:AR052763         +   55.00  173.97  0.1442   1512  ! AR052763 Sequence 99 from pate
gb_pat:E13558           +   55.00  172.81  0.1673   1760  ! E13558 Mutated p53 gene. 6/199
gb_pat:E13737           +   55.00  172.81  0.1673   1760  ! E13737 cDNA encoding human p53
gb_pr7:HUMP53T          +   55.00  172.81  0.1673   1760  ! K03199 Human p53 cellular tumo
gb_pr7:MACP53A          +   55.00  171.17  0.2067   2184  ! L20442 Rhesus monkey p53 mRNA,
gb_sy:AF192534          +   55.00  170.59  0.2225   2355  ! AF192534 Expression vector Ad5
gb_ro:SBU43902          +   52.00  166.88  0.3580    943  ! U43902 Spermophilus beecheyi
gb_ro:MMTSP53           +   52.00  165.20  0.4443   1176  ! AJ001022 Marmata monax mRNA
gb_pr7:U48957           +   50.00  158.03  1.11     1182  ! U48957 Macaca fascicularis
gb_pr5:HS163M094        -   46.00  139.72  11.66    2015  ! AL035288 H.sapiens gene fro

seq_name: gb_pr6:HSP53001

seq_documentation_block:
LOCUS       HSP53001      735 bp    mRNA            PRI       23-JUN-1994
DEFINITION  Human mRNA for mutated p53 transformation suppressor gene.
ACCESSION   X60010
VERSION     X60010.1  GI:506432
KEYWORDS    p53 gene; p53 protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 735)
  AUTHORS   Farrell,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-1991) P.J. Farrell, Ludwig Inst for Cancer Res,
            St Mary's Hospital Med School, Norfolk Place, London W2 1PG, UK
REFERENCE   2  (bases 1 to 735)
  AUTHORS   Farrell,P.J., Allan,G.J., Shanahan,F., Vousden,K.H. and Crook,T.
  TITLE     p53 is frequently mutated in Burkitt's lymphoma cell lines
  JOURNAL   EMBO J. 10 (10), 2879-2887 (1991)
  MEDLINE   92007731
COMMENT     mutated p53 transformation suppressor gene.
FEATURES             Location/Qualifiers
     source          1. .735
                     /organism="Homo sapiens"
                     /isolate="Burkitts lymphoma cell line"
                     /db_xref="taxon:9606"
                     /cell_line="AKATA"
     mRNA            1. .735
                     /gene="p53"
                     /note="cDNA"
                     /evidence=experimental
     gene            1. .735
                     /gene="p53"
     CDS             1. .>735
                     /gene="p53"
                     /note="deletion of a c nucleotide causes a  frameshift at
                     566"
                     /codon_start=1
                     /product="p53  transformation suppressor"
                     /protein_id="CAA42625.1"
                     /db_xref="GI:506433"
                     /db_xref="SPTREMBL:Q15085"
                     /translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLM
                     LSPDDIEQWFTEDPGPDEAPRMPEAAPRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
                     YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAM
                     AIYKQSQHMTEVVRRCPHHERCSDSDGLALLSILSEWKEICVWSIWMTETLFDIVWWC
                     PMSRLRLALTVPPSTTTCVTVPAWAA"
BASE COUNT      154 a    237 c    182 g    162 t
ORIGIN


alignment_scores:
           Quality:   55.00             Length:     11
             Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x HSP53001   ..

 Align seg 1/1  to: HSP53001  from: 1  to: 735

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: gb_pat:A61761
```

```
seq_documentation_block:
LOCUS       A61761       1098 bp    DNA             PAT       09-MAR-1998
DEFINITION  Sequence 1 from Patent WO9711367.
ACCESSION   A61761
VERSION     A61761.1  GI:3715949
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1098)
  AUTHORS   Chene,P. and Hochkeppel,H.
  TITLE     ASSAY FOR IDENTIFYING INHIBITORS OF THE INTERACTION BETWEEN
            PROTEINS p53 AND dm2
  JOURNAL   Patent: WO 9711367-A 1 27-MAR-1997;
            CIBA GEIGY AG (CH)
FEATURES             Location/Qualifiers
     source          1. .1098
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             4. .1092
                     /codon_start=1
                     /product="RESIDUE 1 TO 362 OF HUMAN P53 PROTEIN (NAMED
                     P53D30)"
                     /protein_id="CAA03592.1"
                     /db_xref="GI:3715950"
                     /translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLM
                     LSPDDIEQWFTEDPGPDEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
                     YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAM
                     AIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVV
                     PYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCA
                     CPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRG
                     RERFEMFRELNEALELKDAQAGKEPGGS"
BASE COUNT      248 a    340 c    289 g    221 t
ORIGIN


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x A61761   ..

 Align seg 1/1  to: A61761  from: 1  to: 1098

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      76 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 108

seq_name: gb_pr6:HSP53002

seq_documentation_block:
LOCUS       HSP53002     1179 bp    mRNA            PRI       17-FEB-1997
DEFINITION  Human mRNA for mutated p53 transformation suppressor gene.
ACCESSION   X60011
VERSION     X60011.1  GI:506434
KEYWORDS    p53 gene; p53 protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1179)
  AUTHORS   Farrell,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-1991) P.J. Farrell, Ludwig Inst for Cancer Res,
            St Mary's Hospital Med School, Norfolk Place, London W2 1PG, UK
REFERENCE   2  (bases 1 to 1179)
  AUTHORS   Farrell,P.J., Allan,G.J., Shanahan,F., Vousden,K.H. and Crook,T.
  TITLE     p53 is frequently mutated in Burkitt's lymphoma cell lines
  JOURNAL   EMBO J. 10 (10), 2879-2887 (1991)
  MEDLINE   92007731
COMMENT     mutated p53 transformation suppressor gene.
FEATURES             Location/Qualifiers
     source          1. .1179
                     /organism="Homo sapiens"
                     /isolate="Burkitts lymphoma cell line"
                     /db_xref="taxon:9606"
                     /cell_line="BJAB"
     mRNA            1. .1179
                     /gene="p53"
                     /note="cDNA"
                     /evidence=experimental
     gene            1. .1179
                     /gene="p53"
     CDS             1. .>1179
                     /gene="p53"
                     /codon_start=1
                     /product="p53 transformation  suppressor"
                     /protein_id="CAA42626.1"
                     /db_xref="GI:506435"
                     /db_xref="SPTREMBL:Q16807"
                     /translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLM
                     LSPDDIEQWFTEDPGPDEAPRMPEAAPRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
                     YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAM
                     AIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQRLIRVEGNLRVEYLDDRNTFRHSVVV
                     PYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCA
                     CPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRG
                     RERFEMFRELNEALELKDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDS
                     D"
     variation       578
                     /gene="p53"
                     /note="wild-type sequence"
                     /replace="a"
BASE COUNT      274 a    365 c    307 g    233 t
ORIGIN


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x HSP53002   ..

 Align seg 1/1  to: HSP53002  from: 1  to: 1179

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: gb_pr6:HSP53003

seq_documentation_block:
LOCUS       HSP53003     1179 bp    mRNA            PRI       23-JUN-1994
DEFINITION  Human mRNA for mutated p53 transformation suppressor gene.
ACCESSION   X60012
VERSION     X60012.1  GI:506436
KEYWORDS    p53 gene; p53 protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1179)
  AUTHORS   Farrell,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-1991) P.J. Farrell, Ludwig Inst for Cancer Res,
            St Mary's Hospital Med School, Norfolk Place, London W2 1PG, UK
REFERENCE   2  (bases 1 to 1179)
  AUTHORS   Farrell,P.J., Allan,G.J., Shanahan,F., Vousden,K.H. and Crook,T.
  TITLE     p53 is frequently mutated in Burkitt's lymphoma cell lines
  JOURNAL   EMBO J. 10 (10), 2879-2887 (1991)
  MEDLINE   92007731
COMMENT     mutated p53 transformation suppressor gene.
```

```
FEATURES             Location/Qualifiers
     source          1. .1179
                     /organism="Homo sapiens"
                     /isolate="Burkitts lymphoma cell line"
                     /db_xref="taxon:9606"
                     /cell_line="IARC/BL2"
     mRNA            1. .1179
                     /gene="p53"
                     /note="cDNA"
                     /evidence=experimental
     gene            1. .1179
                     /gene="p53"
     CDS             1. .>1179
                     /gene="p53"
                     /codon_start=1
                     /product="p53 transformation  suppressor"
                     /protein_id="CAA42627.1"
                     /db_xref="GI:506437"
                     /translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLM
                     LSPDDIEQWFTEDPGPDEAPRMPEAAPRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
                     YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAM
                     AIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVV
                     PYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCA
                     CPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRG
                     RERFEMFRELNEALELKDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDS
                     D"
BASE COUNT      275 a    365 c    306 g    233 t
ORIGIN


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x HSP53003   ..

 Align seg 1/1  to: HSP53003  from: 1  to: 1179

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: gb_pr6:HSP53004

seq_documentation_block:
LOCUS       HSP53004     1179 bp    mRNA            PRI       17-FEB-1997
DEFINITION  Human mRNA for mutated p53 transformation suppressor gene.
ACCESSION   X60013
VERSION     X60013.1  GI:506438
KEYWORDS    p53 gene; p53 protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1179)
  AUTHORS   Farrell,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-1991) P.J. Farrell, Ludwig Inst for Cancer Res,
            St Mary's Hospital Med School, Norfolk Place, London W2 1PG, UK
REFERENCE   2  (bases 1 to 1179)
  AUTHORS   Farrell,P.J., Allan,G.J., Shanahan,F., Vousden,K.H. and Crook,T.
  TITLE     p53 is frequently mutated in Burkitt's lymphoma cell lines
  JOURNAL   EMBO J. 10 (10), 2879-2887 (1991)
  MEDLINE   92007731
COMMENT     mutated p53 transformation suppressor gene.
FEATURES             Location/Qualifiers
     source          1. .1179
                     /organism="Homo sapiens"
                     /isolate="Burkitts lymphoma cell line"
                     /db_xref="taxon:9606"
                     /cell_line="IARC/BL30"
     mRNA            1. .1179
                     /gene="p53"
                     /note="cDNA"
                     /evidence=experimental
     gene            1. .1179
                     /gene="p53"
     CDS             1. .>1179
                     /gene="p53"
                     /codon_start=1
                     /product="p53 transformation  suppressor"
                     /protein_id="CAA42628.1"
                     /db_xref="GI:506439"
                     /db_xref="SPTREMBL:Q15086"
                     /translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLM
                     LSPDDIEQWFTEDPGPDEAPRMPEAAPRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
                     YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAM
                     AIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVV
                     PYEPPEVGSDCTTIHYNYMCNSSCMGGTNRRPILTIITLEDSSGNLLGRNSFEVRVCA
                     CPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRG
                     RERFEMFRELNEALELKDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDS
                     D"
     variation       737
                     /gene="p53"
                     /note="wild-type sequence"
                     /replace="t"
BASE COUNT      275 a    366 c    306 g    232 t
ORIGIN


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x HSP53004   ..

 Align seg 1/1  to: HSP53004  from: 1  to: 1179

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: gb_pr6:HSP53005

seq_documentation_block:
LOCUS       HSP53005     1179 bp    mRNA            PRI       17-FEB-1997
DEFINITION  Human mRNA for mutated p53 transformation suppressor gene.
ACCESSION   X60014
VERSION     X60014.1  GI:506440
KEYWORDS    p53 gene; p53 protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1179)
  AUTHORS   Farrell,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-1991) P.J. Farrell, Ludwig Inst for Cancer Res,
            St Mary's Hospital Med School, Norfolk Place, London W2 1PG, UK
REFERENCE   2  (bases 1 to 1179)
  AUTHORS   Farrell,P.J., Allan,G.J., Shanahan,F., Vousden,K.H. and Crook,T.
  TITLE     p53 is frequently mutated in Burkitt's lymphoma cell lines
  JOURNAL   EMBO J. 10 (10), 2879-2887 (1991)
  MEDLINE   92007731
COMMENT     mutated p53 transformation suppressor gene.
FEATURES             Location/Qualifiers
     source          1. .1179
                     /organism="Homo sapiens"
                     /isolate="Burkitts lymphoma cell line"
                     /db_xref="taxon:9606"
                     /cell_line="IARC/BL37"
     mRNA            1. .1179
```

```
                     /gene="p53"
                     /note="cDNA"
                     /evidence=experimental
     gene            1. .1179
                     /gene="p53"
     CDS             1. .>1179
                     /gene="p53"
                     /codon_start=1
                     /product="p53 transformation  suppressor"
                     /protein_id="CAA42629.1"
                     /db_xref="GI:506441"
                     /db_xref="SPTREMBL:Q15087"
                     /translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLM
                     LSPDDIEQWFTEDPGPDEAPRMPEAAPRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
                     YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAM
                     AIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVV
                     PYEPPEVGSDCTTIHYNYICNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCA
                     CPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRG
                     RERFEMFRELNEALELKDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDS
                     D"
     variation       711
                     /gene="p53"
                     /note="wild-type sequence"
                     /replace="g"
BASE COUNT      276 a    365 c    305 g    233 t
ORIGIN


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x HSP53005   ..

 Align seg 1/1  to: HSP53005  from: 1  to: 1179

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: gb_pr6:HSP53006

seq_documentation_block:
LOCUS       HSP53006     1179 bp    mRNA            PRI       17-FEB-1997
DEFINITION  Human mRNA for mutated p53 transformation suppressor gene.
ACCESSION   X60015
VERSION     X60015.1  GI:506442
KEYWORDS    p53 gene; p53 protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1179)
  AUTHORS   Farrell,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-1991) P.J. Farrell, Ludwig Inst for Cancer Res,
            St Mary's Hospital Med School, Norfolk Place, London W2 1PG, UK
REFERENCE   2  (bases 1 to 1179)
  AUTHORS   Farrell,P.J., Allan,G.J., Shanahan,F., Vousden,K.H. and Crook,T.
  TITLE     p53 is frequently mutated in Burkitt's lymphoma cell lines
  JOURNAL   EMBO J. 10 (10), 2879-2887 (1991)
  MEDLINE   92007731
COMMENT     mutated p53 transformation suppressor gene.
FEATURES             Location/Qualifiers
     source          1. .1179
                     /organism="Homo sapiens"
                     /isolate="Burkitts lymphoma cell line"
                     /db_xref="taxon:9606"
                     /cell_line="IARC/BL41"
     mRNA            1. .1179
                     /gene="p53"
                     /note="cDNA"
                     /evidence=experimental
     gene            1. .1179
                     /gene="p53"
     CDS             1. .>1179
                     /gene="p53"
                     /codon_start=1
                     /product="p53 transformation  suppressor"
                     /protein_id="CAA42630.1"
                     /db_xref="GI:506443"
                     /db_xref="SPTREMBL:Q16535"
                     /translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLM
                     LSPDDIEQWFTEDPGPDEAPRMPEAAPRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
                     YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAM
                     AIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVV
                     PYEPPEVGSDCTTIHYNYMCNSSCMGGMNQRPILTIITLEDSSGNLLGRNSFEVRVCA
                     CPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRG
                     RERFEMFRELNEALELKDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDS
                     D"
     variation       743
                     /gene="p53"
                     /note="wild-type sequence"
                     /replace="g"
BASE COUNT      276 a    365 c    305 g    233 t
ORIGIN


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x HSP53006   ..

 Align seg 1/1  to: HSP53006  from: 1  to: 1179

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: gb_pr6:HSP53007

seq_documentation_block:
LOCUS       HSP53007     1179 bp    mRNA            PRI       17-FEB-1997
DEFINITION  Human mRNA for mutated p53 transformation suppressor gene.
ACCESSION   X60016
VERSION     X60016.1  GI:506444
KEYWORDS    p53 gene; p53 protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1179)
  AUTHORS   Farrell,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-1991) P.J. Farrell, Ludwig Inst for Cancer Res,
            St Mary's Hospital Med School, Norfolk Place, London W2 1PG, UK
REFERENCE   2  (bases 1 to 1179)
  AUTHORS   Farrell,P.J., Allan,G.J., Shanahan,F., Vousden,K.H. and Crook,T.
  TITLE     p53 is frequently mutated in Burkitt's lymphoma cell lines
  JOURNAL   EMBO J. 10 (10), 2879-2887 (1991)
  MEDLINE   92007731
COMMENT     mutated p53 transformation suppressor gene.
FEATURES             Location/Qualifiers
     source          1. .1179
                     /organism="Homo sapiens"
                     /isolate="Burkitts lymphoma cell line"
                     /db_xref="taxon:9606"
                     /cell_line="LOUCKES"
     mRNA            1. .1179
                     /gene="p53"
                     /note="cDNA"
```

```
                    /evidence=experimental
     gene           1. .1179
                    /gene="p53"
     CDS            1. .>1179
                    /gene="p53"
                    /codon_start=1
                    /product="p53 transformation  suppressor"
                    /protein_id="CAA42631.1"
                    /db_xref="GI:506445"
                    /db_xref="SPTREMBL:Q15088"
                    /translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLM
                    LSPDDIEQWFTEDPGPDEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
                    YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAM
                    AIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVV
                    PYEPPEVGSDCTTIHYNYMYNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCA
                    CPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRG
                    RERFEMFRELNEALELKDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDS
                    D"
     variation      713
                    /gene="p53"
                    /note="wild-type sequence"
                    /replace="g"
BASE COUNT      276 a    366 c    304 g    233 t
ORIGIN


alignment_scores:
           Quality:   55.00            Length:     11
             Ratio:  5.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x HSP53007   ..

 Align seg 1/1  to: HSP53007  from: 1  to: 1179

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: gb_pr6:HSP53008

seq_documentation_block:
LOCUS       HSP53008     1179 bp    mRNA            PRI       17-FEB-1997
DEFINITION  Human mRNA for mutated p53 transformation suppressor gene.
ACCESSION   X60017
VERSION     X60017.1  GI:506446
KEYWORDS    p53 gene; p53 protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1179)
  AUTHORS   Farrell,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-1991) P.J. Farrell, Ludwig Inst for Cancer Res,
            St Mary's Hospital Med School, Norfolk Place, London W2 1PG, UK
REFERENCE   2  (bases 1 to 1179)
  AUTHORS   Farrell,P.J., Allan,G.J., Shanahan,F., Vousden,K.H. and Crook,T.
  TITLE     p53 is frequently mutated in Burkitt's lymphoma cell lines
  JOURNAL   EMBO J. 10 (10), 2879-2887 (1991)
  MEDLINE   92007731
COMMENT     mutated p53 transformation suppressor gene.
FEATURES             Location/Qualifiers
     source          1. .1179
                     /organism="Homo sapiens"
                     /isolate="Burkitts lymphoma cell line"
                     /db_xref="taxon:9606"
                     /cell_line="NAMALWA"
     mRNA            1. .1179
                     /gene="p53"
                     /note="cDNA"
                     /evidence=experimental
     gene            1. .1179
                     /gene="p53"
     CDS             1. .>1179
                     /gene="p53"
                     /codon_start=1
                     /product="p53 transformation suppressor"
                     /protein_id="CAA42632.1"
                     /db_xref="GI:506447"
                     /db_xref="SPTREMBL:Q16535"
                     /translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLM
                     LSPDDIEQWFTEDPGPDEAPRMPEAAPRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
                     YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAM
                     AIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVV
                     PYEPPEVGSDCTTIHYNYMCNSSCMGGMNQRPILTIITLEDSSGNLLGRNSFEVRVCA
                     CPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRG
                     RERFEMFRELNEALELKDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDS
                     D"
     variation       743
                     /gene="p53"
                     /note="wild-type sequence"
                     /replace="g"
BASE COUNT      276 a    365 c    305 g    233 t
ORIGIN


alignment_scores:
           Quality:   55.00            Length:     11
             Ratio:  5.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x HSP53008   ..

 Align seg 1/1  to: HSP53008  from: 1  to: 1179

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: gb_pr6:HSP53009

seq_documentation_block:
LOCUS       HSP53009     1179 bp    mRNA            PRI       17-FEB-1997
DEFINITION  Human mRNA for mutated p53 transformation suppressor gene.
ACCESSION   X60018
VERSION     X60018.1  GI:506448
KEYWORDS    p53 gene; p53 protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1179)
  AUTHORS   Farrell,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-1991) P.J. Farrell, Ludwig Inst for Cancer Res,
            St Mary's Hospital Med School, Norfolk Place, London W2 1PG, UK
REFERENCE   2  (bases 1 to 1179)
  AUTHORS   Farrell,P.J., Allan,G.J., Shanahan,F., Vousden,K.H. and Crook,T.
  TITLE     p53 is frequently mutated in Burkitt's lymphoma cell lines
  JOURNAL   EMBO J. 10 (10), 2879-2887 (1991)
  MEDLINE   92007731
COMMENT     mutated p53 transformation suppressor gene.
FEATURES             Location/Qualifiers
     source          1. .1179
                     /organism="Homo sapiens"
                     /isolate="Burkitts lymphoma cell line"
                     /db_xref="taxon:9606"
                     /cell_line="P3HR1"
     mRNA            1. .1179
                     /gene="p53"
                     /note="cDNA"
                     /evidence=experimental
     gene            1. .1179
```

```
                     /gene="p53"
     CDS             1. .>1179
                     /gene="p53"
                     /codon_start=1
                     /product="p53 transformation  suppressor"
                     /protein_id="CAA42633.1"
                     /db_xref="GI:506449"
                     /db_xref="SPTREMBL:Q16808"
                     /translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLM
                     LSPDDIEQWFTEDPGPDEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
                     YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAM
                     AIHKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVV
                     PYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCA
                     CPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRG
                     RERFEMFRELNEALELKDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDS
                     D"
     variation       487
                     /gene="p53"
                     /note="wild-type sequence"
                     /replace="t"
BASE COUNT      275 a    367 c    305 g    232 t
ORIGIN


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x HSP53009   ..

 Align seg 1/1  to: HSP53009  from: 1  to: 1179

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: gb_pr6:HSP53010

seq_documentation_block:
LOCUS       HSP53010     1179 bp    mRNA            PRI       17-FEB-1997
DEFINITION  Human mRNA for mutated p53 transformation suppressor gene.
ACCESSION   X60019
VERSION     X60019.1  GI:506450
KEYWORDS    p53 gene; p53 protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1179)
  AUTHORS   Farrell,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-1991) P.J. Farrell, Ludwig Inst for Cancer Res,
            St Mary's Hospital Med School, Norfolk Place, London W2 1PG, UK
REFERENCE   2  (bases 1 to 1179)
  AUTHORS   Farrell,P.J., Allan,G.J., Shanahan,F., Vousden,K.H. and Crook,T.
  TITLE     p53 is frequently mutated in Burkitt's lymphoma cell lines
  JOURNAL   EMBO J. 10 (10), 2879-2887 (1991)
  MEDLINE   92007731
COMMENT     mutated p53 transformation suppressor gene.
FEATURES             Location/Qualifiers
     source          1. .1179
                     /organism="Homo sapiens"
                     /isolate="Burkitts lymphoma cell line"
                     /db_xref="taxon:9606"
                     /cell_line="RAJI"
     mRNA            1. .1179
                     /gene="p53"
                     /note="cDNA"
                     /evidence=experimental
     gene            1. .1179
                     /gene="p53"
     CDS             1. .>1179
                     /gene="p53"
                     /codon_start=1
                     /product="p53 transformation  suppressor"
                     /protein_id="CAA42634.1"
                     /db_xref="GI:506451"
                     /db_xref="SPTREMBL:Q16809"
                     /translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLM
                     LSPDDIEQWFTEDPGPDEAPRMPEAAPRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
                     YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAM
                     AIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFQHSVVV
                     PYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCA
                     CPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRG
                     RERFEMFRELNEALELKDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDS
                     D"
     variation       638
                     /gene="p53"
                     /note="wild-type sequence"
                     /replace="g"
BASE COUNT      276 a    365 c    305 g    233 t
ORIGIN


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x HSP53010   ..

 Align seg 1/1  to: HSP53010  from: 1  to: 1179

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: gb_pr6:HSP53011

seq_documentation_block:
LOCUS       HSP53011     1179 bp    mRNA            PRI       17-FEB-1997
DEFINITION  Human mRNA for mutated p53 transformation suppressor gene.
ACCESSION   X60020
VERSION     X60020.1  GI:506452
KEYWORDS    p53 gene; p53 protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1179)
  AUTHORS   Farrell,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-1991) P.J. Farrell, Ludwig Inst for Cancer Res,
            St Mary's Hospital Med School, Norfolk Place, London W2 1PG, UK
REFERENCE   2  (bases 1 to 1179)
  AUTHORS   Farrell,P.J., Allan,G.J., Shanahan,F., Vousden,K.H. and Crook,T.
  TITLE     p53 is frequently mutated in Burkitt's lymphoma cell lines
  JOURNAL   EMBO J. 10 (10), 2879-2887 (1991)
  MEDLINE   92007731
COMMENT     mutated p53 transformation suppressor gene.
FEATURES             Location/Qualifiers
     source          1. .1179
                     /organism="Homo sapiens"
                     /isolate="Burkitts lymphoma cell line"
                     /db_xref="taxon:9606"
                     /cell_line="RAMOS"
     mRNA            1. .1179
                     /gene="p53"
                     /note="cDNA"
                     /evidence=experimental
     gene            1. .1179
                     /gene="p53"
     CDS             1. .>1179
```

```
                    /gene="p53"
                    /codon_start=1
                    /product="p53 transformation  suppressor"
                    /protein_id="CAA42635.1"
                    /db_xref="GI:506453"
                    /db_xref="SPTREMBL:Q16810"
                    /translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLM
                    LSPDDIEQWFTEDPGPDEAPRMPEAAPRVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
                    YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAM
                    AIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVV
                    PYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTDITLEDSSGNLLGRNSFEVRVCA
                    CPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRG
                    RERFEMFRELNEALELKDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDS
                    D"
     variation      760
                    /gene="p53"
                    /note="wild-type sequence"
                    /replace="a"
     variation      761
                    /gene="p53"
                    /note="wild-type sequence"
                    /replace="t"
BASE COUNT      275 a    365 c    307 g    232 t
ORIGIN


alignment_scores:
           Quality:   55.00             Length:     11
             Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x HSP53011   ..

 Align seg 1/1  to: HSP53011  from: 1  to: 1179

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: gb_pat:A61360

seq_documentation_block:
LOCUS       A61360       1182 bp    DNA             PAT       09-MAR-1998
DEFINITION  Sequence 2 from Patent WO9709343.
ACCESSION   A61360
VERSION     A61360.1  GI:3715771
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1182)
  AUTHORS   Tocque,B., Dubs-Poterszman,M. and Wasylyk,B.
  TITLE     ANTAGONISTS OF THE ONCOGENIC ACTIVITY OF THE PROTEIN MDM2, AND USE
            THEREOF IN THE TREATMENT OF CANCERS
  JOURNAL   Patent: WO 9709343-A 2 13-MAR-1997;
            RHONE POULENC RORER SA (FR)
COMMENT     Other publication FR 2738151 970307.
FEATURES             Location/Qualifiers
     source          1. .1182
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .1182
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAA03589.1"
                     /db_xref="GI:3715772"
                     /translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLM
                     LSPDDIEQWFTEDPGPDEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
                     YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAM
                     AIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVV
                     PYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCA
                     CPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRG
                     RERFEMFRELNEALELKDAQAGKEPGGSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDS
                     D"
BASE COUNT      276 a    366 c    306 g    234 t
ORIGIN


alignment_scores:
           Quality:   55.00             Length:     11
             Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x A61360   ..

 Align seg 1/1  to: A61360  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: gb_pat:AR052878

seq_documentation_block:
LOCUS       AR052878     1182 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 215 from patent US 5833975.
ACCESSION   AR052878
VERSION     AR052878.1  GI:5977740
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1182)
  AUTHORS   Paoletti,E., Tartaglia,J. and Cox,W.I.
  TITLE     Canarypox virus expressing cytokine and/or tumor-associated antigen
            DNA sequence
  JOURNAL   Patent: US 5833975-A 215 10-NOV-1998;
FEATURES             Location/Qualifiers
     source          1. .1182
                     /organism="unknown"
BASE COUNT      276 a    365 c    307 g    234 t
ORIGIN


alignment_scores:
           Quality:   55.00             Length:     11
             Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x AR052878   ..

 Align seg 1/1  to: AR052878  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: gb_pat:AR061772

seq_documentation_block:
LOCUS       AR061772     1182 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 92 from patent US 5843654.
ACCESSION   AR061772
VERSION     AR061772.1  GI:5989463
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1182)
  AUTHORS   Heisler,L.M., Fors,L. and Brow,M.AnnD.
  TITLE     Rapid detection of mutations in the p53 gene
  JOURNAL   Patent: US 5843654-A 92 01-DEC-1998;
FEATURES             Location/Qualifiers
```

```
     source          1. .1182
                     /organism="unknown"
BASE COUNT      276 a    366 c    306 g    234 t
ORIGIN


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x AR061772   ..

 Align seg 1/1  to: AR061772  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105
```

```
OM of: US-08-860-232-1 to: N_Geneseq_36:*   out_format : pfs

Date: Dec 12, 2000  5:51 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US08860232/runat_11122000_154523_20768/app_query.fasta_1.67
 -DB=N_Geneseq_36 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=2000000000
 -USER=US08860232_@CGN1_1_108 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-08-860-232-1
Query length: 11
Database: N_Geneseq_36:*
Database sequences: 480022
Database length: 187831343
Search time (sec): 71.020000



score_list:
Sequence           Strd Orig    ZScore    EScore Len  ! Documentation  ..
/SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:Q62357 +   55.00  175.33  0.0561    723  ! Fr
/SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:Z52304 +   55.00  171.74  0.0889   1070  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1995.DAT:Q97854 +   55.00  170.83  0.0999   1181  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:Z27570 +   55.00  170.83  0.0999   1181  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1992.DAT:Q22995 +   55.00  170.82  0.1000   1182  ! S
/SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:Q67884 +   55.00  170.82  0.1000   1182  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T29719 +   55.00  170.82  0.1000   1182  ! W
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:Z08529 +   55.00  170.82  0.1000   1182  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T27665 +   55.00  170.80  0.1003   1185  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T27663 +   55.00  170.66  0.1021   1203  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T27664 +   55.00  170.66  0.1022   1204  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T32836 +   55.00  170.57  0.1033   1215  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V21730 +   55.00  170.37  0.1060   1242  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X04533 +   55.00  169.93  0.1121   1303  ! D
/SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:D00088 +   55.00  169.90  0.1125   1307  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T32831 +   55.00  169.83  0.1135   1317  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V44695 +   55.00  169.83  0.1135   1317  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V18151 +   55.00  169.83  0.1135   1317  ! N
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V21414 +   55.00  169.83  0.1135   1317  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:Q67864 +   55.00  168.74  0.1306   1483  ! V
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:Z08435 +   55.00  168.74  0.1307   1484  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:Q67863 +   55.00  168.57  0.1336   1512  ! V
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:Z08434 +   55.00  168.57  0.1336   1512  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:T72724 +   55.00  167.99  0.1438   1610  ! p
/SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:T79880 +   55.00  167.17  0.1597   1760  ! D
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:X75765 +   55.00  167.17  0.1597   1760  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:Q62365 +   44.00  158.06  0.5138     75  ! Fra
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V00615 +   44.00  136.80  7.85      761  ! Ly
/SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T09162 -   40.00  142.45  3.81       91  ! Myo
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V22760 -   40.00  142.45  3.81       91  ! Amp
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V20710 -   40.00  142.45  3.81       91  ! Myo
/SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T09163 -   40.00  138.17  6.59      145  ! My
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V30458 +   40.00   62.82  9.8e+04  534720  !
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V30459 +   40.00   62.79  9.9e+04  536165  !
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:V86679 -   39.00  124.32  38.92     450  ! ES
/SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:Z95203 +   39.00  111.49  201.76   1822  ! H
/SIDS6/gcgdata/geneseq/geneseqn/NA1992.DAT:Q27944 +   38.50  110.18  238.60   1740  ! M
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V30305 +   38.00  113.69  152.13    983  ! Ba
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V30302 +   38.00  108.99  278.04   1641  ! E
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:Z40490 -   38.00  106.10  402.89   2249  ! M
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X23521 +   38.00   87.76  4.2e+03  16595
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:Z23904 +   38.00   77.64  1.5e+04  49999
/SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:X83003 -   38.00   72.52  2.9e+04  87359
/SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:T58840 +   38.00   55.15  2.4e+05  580073
/SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T42063 -   38.00   44.61  6.4e+05  183012

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:Q62357

seq_documentation_block:
ID   Q62357 standard; cDNA; 723 BP.
XX
AC   Q62357;
XX
DT   18-NOV-1994  (first entry)
XX
DE   Fragment coding for human p53 amino acids 1-241.
XX
KW   Human nuclear phosphoprotein p53; tumour suppressor gene product;
KW   anti-oncogene; cancer; tumour; antibody binding region; epitope; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..723
FT                   /*tag= a
FT                   /product= p53(1-241)
FT                   /note= "partial CDS, therefore sequence does not
FT                          contain stop codon"
XX
PN   WO9408241-A.
XX
PD   14-APR-1994.
XX
PF   30-SEP-1993;  93WO-EP02666.
XX
PR   30-SEP-1992;  92DE-4232823.
XX
PA   (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM STIFTUNG.
XX
PI   Klein R,  Schranz P,  Tessmer C,  Volkmann M,  Zentgraf H;
XX
DR   WPI; 1994-135732/16.
DR   P-PSDB; R51872.
XX
PT   Non-radioactive detection of p53 specific antibodies - by capture
PT   on immobilised p53 or its fragments, then reaction with labelled
PT   second antibody, for diagnosis of tumours and suitable for
PT   screening
XX
PS   Claim 13; Page 20; 35pp; German.
XX
CC   Antibodies specific for p53 are detected by binding to immobilised
CC   fragments of the p53 gene product containing the antibody-binding
CC   region. Preferred fragments contain amino acids 1-241, 40-349,
CC   40-393, 66-241, 66-393, 237-349, 237-393 and esp. 9-33, 37-52 or
CC   368-386. See Q62357-Q62366 for nucleic acid sequences coding for
CC   each of these fragments.
XX
SQ   Sequence 723 BP; 152 A; 236 C; 175 G; 160 T; 0 other;


alignment_scores:
           Quality:   55.00          Length:     11
             Ratio:  5.000            Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x Q62357   ..

 Align seg 1/1  to: Q62357  from: 1  to: 723

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
```

```
     73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:Z52304

seq_documentation_block:
ID   Z52304 standard; cDNA; 1070 BP.
XX
AC   Z52304;
XX
DT   24-JUL-2000  (first entry)
XX
DE   Human p35 (p53 isoform) cDNA.
XX
KW   Human p53 isoform; p35; marker; hypoxia; myocardial infarction;
KW   cell proliferation; cytostatic; proliferative; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             136..921
FT                   /*tag= a
FT                   /product= "Human p35 protein"
XX
PN   WO200022127-A1.
XX
PD   20-APR-2000.
XX
PF   07-OCT-1999;   99WO-US23319.
XX
PR   09-OCT-1998;   98US-0103849.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Dell'acqua G,  Mann MJ,  Dzau VJ;
XX
DR   WPI; 2000-317984/27.
DR   P-PSDB; Y70714.
XX
PT   Novel isoform of p53 useful as a marker of myocardial infarction and
PT   for controlling cellular proliferation  -
XX
PS   Claim 18; Fig 4; 25pp; English.
XX
CC   The present cDNA sequence encodes the human p53 isoform, p35. This p53
CC   isoform is truncated to eliminate a substantial portion of the C-terminal
CC   end of the protein. The deleted portion correspond to those encoded by
CC   exon 7 and by exons corresponding to amino acids lying C-terminal to
CC   exon 7. p35 is useful as a marker of myocardial infarction (by the
CC   indication of hypoxia) and in the control of cell proliferation both
CC   in vivo and in vitro. It was able to increase the transactivation
CC   achieved through co-transfection with wild type p53. The combination of
CC   p35 with wild type p53 produce an enhanced inhibition of primary
CC   incorporation at 24 hours compared to wild type p53 alone.
XX
SQ   Sequence 1070 BP; 233 A; 338 C; 282 G; 217 T; 0 other;


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x Z52304    ..

 Align seg 1/1  to: Z52304  from: 1  to: 1070

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
     208 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 240

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1995.DAT:Q97854

seq_documentation_block:
ID   Q97854 standard; cDNA; 1181 BP.
XX
AC   Q97854;
XX
DT   06-DEC-1995  (first entry)
XX
DE   Human p53 cDNA.
XX
KW   Ubiquitin-conjugating enzyme; p53 protein; cell cycle;
KW   cell proliferation; cancer; psoriasis; fibrosis; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1181
FT                   /*tag= a
XX
PN   WO9518974-A.
XX
PD   13-JUL-1995.
XX
PF   04-JAN-1995;   95WO-US00164.
XX
PR   13-SEP-1994;   94US-0305520.
PR   04-JAN-1994;   94US-0176937.
PR   23-MAY-1994;   94US-0247904.
PR   27-MAY-1994;   94US-0250795.
XX
PA   (MITO-) MITOTIX INC.
XX
PI   Cottarel G,  Draetta G,  Eckstein JW,  Gyuris J,  Rolfe M;
XX
DR   WPI; 1995-255137/33.
DR   P-PSDB; R79658.
XX
PT   Identifying inhibitors of ubiquitin mediated proteolysis of cell cycle
PT   regulatory proteins - also new ubiquitin conjugating enzymes, their
PT   related nucleic acid, vectors, antibodies etc., useful for regulating
PT   e.g. cell proliferation
XX
PS   Disclosure; Page 105-106; 157pp; English.
XX
CC   Human p53 cDNA (given in Q97854) was amplified from a HeLa cell
CC   cDNA library using the primers given in Q97852-53.  The gene
CC   was subcloned into a baculovirus vector for expression of
CC   recombinant p53 in Sf9 insect cells for use as a component of
CC   an in vitro ubiquitin conjugating system.
XX
SQ   Sequence 1181 BP; 275 A; 366 C; 306 G; 234 T; 0 other;


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x Q97854    ..

 Align seg 1/1  to: Q97854  from: 1  to: 1181

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:Z27570

seq_documentation_block:
ID   Z27570 standard; cDNA; 1181 BP.
XX
AC   Z27570;
XX
```

```
DT   15-DEC-1999  (first entry)
XX
DE   Human p53 coding sequence.
XX
KW   Ubiquitin conjugating enzyme; UbCE; ubiquitin-mediated proteolysis;
KW   cell-cycle regulatory protein; ubiquitination inhibitor; atherosclerosis;
KW   proliferative disorder; cancer; restenosis; tissue connective disorder;
KW   wound healing; fibrosis disorder; rheumatoid arthritis; scleroderma;
KW   insulin dependent diabetes mellitus; glomerulonephritis; cirrhosis;
KW   diagnosis; therapy; p53; ds.
XX
OS   Homo sapiens.
XX
PN   US5968761-A.
XX
PD   19-OCT-1999.
XX
PF   07-JUN-1995;   95US-0486663.
XX
PR   04-JAN-1994;   94US-0176937.
PR   23-MAY-1994;   94US-0247904.
PR   27-MAY-1994;   94US-0250795.
PR   13-SEP-1994;   94US-0305520.
XX
PA   (MITO-) MITOTIX INC.
XX
PI   Chiu MI,  Cottarel G,  Berlin V,  Damagnez V,  Draetta G,  Rolfe M;
XX
DR   WPI; 1999-590402/50.
DR   P-PSDB; Y39970.
XX
PT   Identifying ubiquitination inhibitors using novel ubiquitin conjugating
PT   enzymes -
XX
PS   Example 2; Column 97-100; 61pp; English.
XX
CC   This sequence encodes the human p53 protein. The invention relates to
CC   assays for identifying an inhibitor of ubiquitin-mediated proteolysis of
CC   a cell-cycle regulatory protein comprising contacting a candidate agent
CC   with an ubiquitin-conjugating system and measuring the level of
CC   ubiquitination. The ubiquitin-conjugating system comprises:
CC   (a) a reconstituted protein mixture including a ubiquitin conjugating
CC   enzyme (UbCE) produced by the expression of a nucleic acid which
CC   hybridizes under high stringency conditions to human UbCE, Candida
CC   albicans UbCE, or Schizosaccharomyces pombe UbCE coding sequences;
CC   (b) a regulatory protein; and (c) ubiquitin. The polynucleotides are
CC   useful for identifying ubiquitination inhibitors. The polynucleotides,
CC   polypeptides, antisense compounds and antibodies against them may also be
CC   useful for the treatment and/or diagnosis of proliferative disorders
CC   (e.g. cancer, atherosclerosis, or restenosis), tissue connective
CC   disorders, controlling wound healing, and disorders characterized by
CC   fibrosis (e.g. rheumatoid arthritis, insulin dependent diabetes mellitus,
CC   glomerulonephritis, cirrhosis, and scleroderma).
XX
SQ   Sequence 1181 BP; 275 A; 366 C; 306 G; 234 T; 0 other;


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x Z27570   ..

 Align seg 1/1  to: Z27570  from: 1  to: 1181

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1992.DAT:Q22995

seq_documentation_block:
ID   Q22995 standard; DNA; 1182 BP.
XX
AC   Q22995;
XX
DT   23-JUL-1992  (first entry)
XX
DE   Sequence encoding 53 kD cellular protein.
XX
KW   Cancer therapy; cancer suppressor gene; oncogenesis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1182
FT                   /*tag=  a
XX
PN   EP475623-A.
XX
PD   18-MAR-1992.
XX
PF   23-AUG-1991;   91EP-0307791.
XX
PR   24-AUG-1990;   90US-0573405.
XX
PA   (REGC ) UNIV OF CALIFORNIA.
XX
PI   Lee WH,  Chen PL;
XX
DR   WPI; 1992-090221/12.
DR   P-PSDB; R22238.
XX
PT   Cloned p53 cDNA and protein prods. - for suppression of
PT   neoplastic phenotype e.g. in osteo-sarcuma(s), leukaemia(s),
PT   lymphoma(s), etc.
XX
PS   Claim 5; Page 15; 25pp; English.
XX
CC   p53 cDNA, or its gene prods., can be used to suppress and eradicate
CC   cancers caused by defective, mutant or absent cancer suppressor
CC   genes.  Variant forms of p53 are found in human breast, lung or
CC   colon carcinoma, lymphoma, leukaemia, etc., suggesting that mutation
CC   of the p53 genes is involved in oncogenesis.  Specifically 273 Arg
CC   is replaced by 273 His, a mutation found exclusively in tumour cells.
XX
SQ   Sequence 1182 BP; 277 A; 368 C; 303 G; 234 T; 0 other;


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x Q22995   ..

 Align seg 1/1  to: Q22995  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:Q67884

seq_documentation_block:
ID   Q67884 standard; DNA; 1182 BP.
XX
AC   Q67884;
XX
DT   23-MAR-1995  (first entry)
XX
DE   Human p53 DNA.
XX
```

```
KW   Polymerase chain reaction; primer; amplify; NYVAC; ALVAC; recombinant;
KW   murine; interleukin-2; IL-2; pRW825; pmut-1; pBS-SK; pMM151; TK vector;
KW   plasmid; vaccinia; H6 promoter; amplify; primer; antigenic response;
KW   polymerase chain reaction; poxvirus; pSD542; immunological response;
KW   pathogen; human; interferon; IFN; ss.
XX
OS   Synthetic.
XX
PN   WO9416716-A.
XX
PD   04-AUG-1994.
XX
PF   21-JAN-1994;   94WO-US00888.
XX
PR   21-JAN-1993;   93US-0007115.
PR   19-JAN-1994;   94US-0184009.
XX
PA   (VIRO-) VIROGENETICS CORP.
XX
PI   Cox WI,  Paoletti E,  Tartaglia J;
XX
DR   WPI; 1994-263767/32.
XX
PT   Attenuated recombinant virus used for cancer therapy - comprises
PT   DNA encoding cytokine and/or tumour associated antigen
XX
PS   Example 32; Fig 39; 232pp; English.
XX
CC   This sequence represents the wildtype human p53 gene from the translation
CC   initiation codon to the stop codon.  This sequence was used in the
CC   construction of an ALVAC-based recombinant virus containing a mutant
CC   form of the human p53 gene.  The mutant form has a G>A substitution at
CC   position 524, changing an Arg residue at position 175 to a His residue.
CC   The plasmid pMM110 (see also Q67864) contains the vaccinia H6 promoter
CC   and the wild type human p53 gene in the ALVAC C5 insertion site.  The
CC   mutant p53 gene was obtained from plasmid Cx22A and cloned into pMM110
CC   to generate pMM143.  Recombination between pMM143 and ALVAC rescuing
CC   virus produced recombinant virus vCP270, which contains the vaccinia H6
CC   promoted mutated human p53 in the C5 locus.  The resulting virus may be
CC   used in a composition for inducing an antigenic or immunological
CC   response, ie. for immunisation against pathogens.
XX
SQ   Sequence 1182 BP; 276 A; 365 C; 307 G; 234 T; 0 other;


alignment_scores:
           Quality:   55.00             Length:     11
             Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x Q67884   ..

 Align seg 1/1  to: Q67884  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T29719

seq_documentation_block:
ID   T29719 standard; cDNA; 1182 BP.
XX
AC   T29719;
XX
DT   29-OCT-1996  (first entry)
XX
DE   Wild type p53 gene sequence.
XX
KW   p53 gene; cancer; carcinoma; neoplastic; neoplasia; phenotype;
KW   osteosarcoma cells; lung carcinoma cells; lymphoma cells;
KW   leukaemia cells; soft tissue sarcoma cells; breast cells;
KW   bladder cells; prostate carcinoma cell; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1182
FT                   /*tag= a
FT                   /product= p53 protein.
FT   misc_difference 19..21
FT                   /*tag= b
FT                   /transl_except= CAT encodes Aspartic acid.
XX
PN   EP710722-A1.
XX
PD   08-MAY-1996.
XX
PF   23-AUG-1991;   91EP-0307791.
XX
PR   24-AUG-1990;   90US-0573405.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Chen P,  Lee W;
XX
DR   WPI; 1996-223439/23.
DR   P-PSDB; R91933.
XX
PT   Use of wild-type p53 gene - in a medicament for suppressing the
PT   neoplastic phenotype of a cancer cell lacking wild-type p53 protein
XX
PS   Claim 1; Page 5; 25pp; English.
XX
CC   The wild-type p53 gene can be used in the production of a medicament
CC   for suppressing the neoplastic phenotype of a cancer cell lacking
CC   endogenous wild type p53 protein. Cancer cells suppressed in such
CC   fashion include osteosarcoma cells, lung carcinoma cells, lymphoma
CC   cells, leukaemia cells, soft tissue sarcoma cells or breast, bladder
CC   or prostate carcinoma cells.
XX
SQ   Sequence 1182 BP; 278 A; 366 C; 304 G; 234 T; 0 other;


alignment_scores:
           Quality:   55.00             Length:     11
             Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x T29719   ..

 Align seg 1/1  to: T29719  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:Z08529

seq_documentation_block:
ID   Z08529 standard; DNA; 1182 BP.
XX
AC   Z08529;
XX
DT   19-OCT-1999  (first entry)
XX
DE   Human p53 gene.
XX
KW   Attenuated recombinant virus; cytokine; tumour associated antigen;
KW   NYVAC recombinant virus; ALVAC recombinant virus; gene therapy; rabies;
KW   cancer; tumour necrosis factor; nuclear phosphoprotein; p53; IL-2; GMCSF;
KW   interleukin; interferon; IFN-gamma; IL-4; melanoma associated antigen;
KW   carcinoembryonic antigen; immunisation; antigenic; poxvirus; influenza;
KW   immunological response; immunotherapy; vaccine; Newcastle Disease; ss.
```

```
XX
OS   Homo sapiens.
XX
PN   US5942235-A.
XX
PD   24-AUG-1999.
XX
PF   02-JUN-1995;   95US-0458356.
XX
PR   02-JUN-1995;   95US-0458356.
PR   24-DEC-1981;   81US-0334456.
PR   08-DEC-1982;   82US-0446824.
PR   19-JUN-1984;   84US-0622135.
PR   27-AUG-1987;   87US-0090209.
PR   28-AUG-1987;   87US-0090711.
PR   20-OCT-1987;   87US-0110335.
PR   25-APR-1988;   88US-0186054.
PR   23-AUG-1988;   88US-0234390.
PR   08-MAR-1989;   89US-0320471.
PR   14-FEB-1990;   90US-0478179.
PR   14-JUN-1990;   90US-0537882.
PR   14-JUN-1990;   90US-0537890.
PR   07-JAN-1991;   91US-0638080.
PR   07-MAR-1991;   91US-0666056.
PR   11-JUN-1991;   91US-0713967.
PR   16-DEC-1991;   91US-0805567.
PR   03-MAR-1992;   92US-0847977.
PR   06-MAR-1992;   92US-0847951.
PR   04-MAY-1992;   92US-0881995.
PR   22-JUL-1992;   92US-0918278.
PR   20-JAN-1993;   93US-0007115.
PR   19-JAN-1994;   94US-0184009.
PR   14-APR-1994;   94US-0228926.
PR   13-SEP-1994;   94US-0306259.
XX
PA   (HEAL-) HEALTH RES INC.
XX
PI   Paoletti E;
XX
DR   WPI; 1999-493494/41.
XX
PT   Recombinant poxviruses comprising exogenous DNA encoding antigenic
PT   determinants useful in immunotherapy to immunize against cancers and
PT   other diseases such as influenza, Newcastle Disease and rabies
XX
PS   Example 32; Fig 39; 163pp; English.
XX
CC   The present invention describes a recombinant poxvirus (I), comprising
CC   exogenous DNA encoding an antigenic determinant of a pathogen which is
CC   then expressed in vivo in infected host cells after administration to a
CC   patient and therefore induces an immunological response. (I) may be used
CC   to vaccinate patients against a wide range of diseases and disorders
CC   depending on the type of antigen encoded by the exogenous DNA. (I) may
CC   be used to vaccinate against diseases such as rabies, influenza and
CC   Newcastle Disease. It is particularly useful for immunising against
CC   cancers. The poxvirus (I) also provides a means of manipulating
CC   lymphocytes and tumour cells for use in cell-based immunotherapeutic
CC   modalities for cancer. (I) also have enhanced safety compared to
CC   unattenuated viruses (attenuation reduces the virulence of the viruses)
CC   and known recombinant poxvirus vaccines. This increased level of safety
CC   reduces the possibility of a 'runaway' infection in the host and reduces
CC   the chance of transmission from vaccinated to unvaccinated individuals
CC   and contamination of the environment. The present sequence represents a
CC   human p53 gene used in the exemplification of the present invention.
XX
SQ   Sequence 1182 BP; 276 A; 365 C; 307 G; 234 T; 0 other;


alignment_scores:
           Quality:   55.00            Length:     11
             Ratio:  5.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000
```

```
alignment_block:
 US-08-860-232-1 x Z08529   ..

 Align seg 1/1  to: Z08529  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T27665

seq_documentation_block:
ID   T27665 standard; DNA; 1185 BP.
XX
AC   T27665;
XX
DT   14-NOV-1996  (first entry)
XX
DE   Human p53 gene sequence.
XX
KW   p53; mutant; mutation; cleavage; nuclease; cleavase; Thermus;
KW   Escherichia; Saccharomyces; Campylobacter; Mycobacterium; Shigella;
KW   Staphylococcus; identification; detection; ds.
XX
OS   Homo sapiens.
XX
PN   WO9615267-A1.
XX
PD   23-MAY-1996.
XX
PF   09-NOV-1995;   95WO-US14673.
XX
PR   30-AUG-1995;   95US-0520946.
PR   09-NOV-1994;   94US-0337164.
PR   09-MAR-1995;   95US-0402601.
PR   07-JUN-1995;   95US-0484956.
XX
PA   (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI   Brow MAD,  Dahlberg JE,  Fors L,  Heisler LM,  Lyamichev VI;
PI   Oldenburg MC, Olive DM;
XX
DR   WPI; 1996-259862/26.
XX
PT   Cleavage of nucleic acids to detect mutation(s) - allows detection
PT   esp. in human p53 gene, to identify strains of microorganisms and
PT   viruses
XX
PS   Claim 28; Page 291; 433pp; English.
XX
CC   Cleavage of nucleic acids using an enzyme, especially a nuclease
CC   selected from the group consisting of Cleavase (RTM) BN enzyme,
CC   Thermus aquaticus DNA polymerase, Thermus thermophilus DNA
CC   polymerase, Escherichia coli ExoIII and the Saccharomyces cerevisiae
CC   Rad1/Rad10 complex.  The nucleic acid substrate is preferably an
CC   oligonucleotide containing a human p53 gene sequence or
CC   alternatively, microbial gene sequences.  Cleavage products are
CC   compared to the cleavage products of reference gene sequences.  The
CC   method is used for detecting mutation in the human p53 gene; for
CC   identifying strains of microorganisms, especially bacteria selected
CC   from the the group of members of the genera Campylobacter,
CC   Escherichia, Mycobacterium, Salmonella, Shigella and Staphylococcus.
CC   The method may also be used for the identification of viruses,
CC   especially hepatitis C virus and simian immunodeficiency virus..
XX
SQ   Sequence 1185 BP; 277 A; 366 C; 305 G; 235 T; 2 other;


alignment_scores:
           Quality:   55.00            Length:     11
             Ratio:  5.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000
```

```
alignment_block:
 US-08-860-232-1 x T27665   ..

 Align seg 1/1  to: T27665  from: 1  to: 1185

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T27663

seq_documentation_block:
ID   T27663 standard; DNA; 1203 BP.
XX
AC   T27663;
XX
DT   14-NOV-1996  (first entry)
XX
DE   Human p53 gene sequence.
XX
KW   p53; mutant; mutation; cleavage; nuclease; cleavase; Thermus;
KW   Escherichia; Saccharomyces; Campylobacter; Mycobacterium; Shigella;
KW   Staphylococcus; identification; detection; ds.
XX
OS   Homo sapiens.
XX
PN   WO9615267-A1.
XX
PD   23-MAY-1996.
XX
PF   09-NOV-1995;   95WO-US14673.
XX
PR   30-AUG-1995;   95US-0520946.
PR   09-NOV-1994;   94US-0337164.
PR   09-MAR-1995;   95US-0402601.
PR   07-JUN-1995;   95US-0484956.
XX
PA   (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI   Brow MAD,  Dahlberg JE,  Fors L,  Heisler LM,  Lyamichev VI;
PI   Oldenburg MC, Olive DM;
XX
DR   WPI; 1996-259862/26.
XX
PT   Cleavage of nucleic acids to detect mutation(s) - allows detection
PT   esp. in human p53 gene, to identify strains of microorganisms and
PT   viruses
XX
PS   Claim 28; Page 289-290; 433pp; English.
XX
CC   Cleavage of nucleic acids using an enzyme, especially a nuclease
CC   selected from the group consisting of Cleavase (RTM) BN enzyme,
CC   Thermus aquaticus DNA polymerase, Thermus thermophilus DNA
CC   polymerase, Escherichia coli ExoIII and the Saccharomyces cerevisiae
CC   Rad1/Rad10 complex.  The nucleic acid substrate is preferably an
CC   oligonucleotide containing a human p53 gene sequence or
CC   alternatively, microbial gene sequences.  Cleavage products are
CC   compared to the cleavage products of reference gene sequences.  The
CC   method is used for detecting mutation in the human p53 gene; for
CC   identifying strains of microorganisms, especially bacteria selected
CC   from the the group of members of the genera Campylobacter,
CC   Escherichia, Mycobacterium, Salmonella, Shigella and Staphylococcus.
CC   The method may also be used for the identification of viruses,
CC   especially hepatitis C virus and simian immunodeficiency virus.
XX
SQ   Sequence 1203 BP; 277 A; 366 C; 306 G; 235 T; 19 other;


alignment_scores:
           Quality:   55.00            Length:     11
             Ratio:  5.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000
```

```
alignment_block:
 US-08-860-232-1 x T27663   ..

 Align seg 1/1  to: T27663  from: 1  to: 1203

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T27664

seq_documentation_block:
ID   T27664 standard; DNA; 1204 BP.
XX
AC   T27664;
XX
DT   14-NOV-1996  (first entry)
XX
DE   Human p53 gene sequence.
XX
KW   p53; mutant; mutation; cleavage; nuclease; cleavase; Thermus;
KW   Escherichia; Saccharomyces; Campylobacter; Mycobacterium; Shigella;
KW   Staphylococcus; identification; detection; ds.
XX
OS   Homo sapiens.
XX
PN   WO9615267-A1.
XX
PD   23-MAY-1996.
XX
PF   09-NOV-1995;   95WO-US14673.
XX
PR   30-AUG-1995;   95US-0520946.
PR   09-NOV-1994;   94US-0337164.
PR   09-MAR-1995;   95US-0402601.
PR   07-JUN-1995;   95US-0484956.
XX
PA   (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI   Brow MAD,  Dahlberg JE,  Fors L,  Heisler LM,  Lyamichev VI;
PI   Oldenburg MC, Olive DM;
XX
DR   WPI; 1996-259862/26.
XX
PT   Cleavage of nucleic acids to detect mutation(s) - allows detection
PT   esp. in human p53 gene, to identify strains of microorganisms and
PT   viruses
XX
PS   Claim 28; Page 290-291; 433pp; English.
XX
CC   Cleavage of nucleic acids using an enzyme, especially a nuclease
CC   selected from the group consisting of Cleavase (RTM) BN enzyme,
CC   Thermus aquaticus DNA polymerase, Thermus thermophilus DNA
CC   polymerase, Escherichia coli ExoIII and the Saccharomyces cerevisiae
CC   Rad1/Rad10 complex.  The nucleic acid substrate is preferably an
CC   oligonucleotide containing a human p53 gene sequence or
CC   alternatively, microbial gene sequences.  Cleavage products are
CC   compared to the cleavage products of reference gene sequences.  The
CC   method is used for detecting mutation in the human p53 gene; for
CC   identifying strains of microorganisms, especially bacteria selected
CC   from the the group of members of the genera Campylobacter,
CC   Escherichia, Mycobacterium, Salmonella, Shigella and Staphylococcus.
CC   The method may also be used for the identification of viruses,
CC   especially hepatitis C virus and simian immunodeficiency virus.
XX
SQ   Sequence 1204 BP; 277 A; 367 C; 306 G; 234 T; 20 other;


alignment_scores:
           Quality:   55.00            Length:     11
             Ratio:  5.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000
```

```
alignment_block:
 US-08-860-232-1 x T27664    ..

 Align seg 1/1  to: T27664  from: 1  to: 1204

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T32836

seq_documentation_block:
ID   T32836 standard; DNA; 1215 BP.
XX
AC   T32836;
XX
DT   06-NOV-1996  (first entry)
XX
DE   Human p53 EcoRI-SalI fragment.
XX
KW   p53 protein; tumour suppressor; tetramerisation domain;
KW   chimaeric protein; gene therapy; vector; cell proliferation; cancer;
KW   apoptosis; autoimmune disease; immune tolerance; pGEMhump53; ds.
XX
OS   Chimeric Homo sapiens;
OS   Chimeric synthetic.
XX
PN   WO9616989-A1.
XX
PD   06-JUN-1996.
XX
PF   27-NOV-1995;   95WO-US15353.
XX
PR   01-JUN-1995;   95US-0456623.
PR   28-NOV-1994;   94US-0347792.
PR   28-APR-1995;   95US-0431357.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Halazonetis TD;
XX
DR   WPI; 1996-286828/29.
XX
PT   New chimaeric p53 protein with heterologous tetramerisation domain
PT   - and related DNA and vectors, useful for treating abnormal cell
PT   proliferation, esp. cancer, auto-immune disease, etc.
XX
PS   Example 1; Page 85-86; 123pp; English.
XX
CC   An EcoRI-SalI DNA fragment (T32836) of plasmid pGEMhump53wt
CC   comprises a coding sequence (see also T32831) for human wild-type
CC   tumour suppressor p53 (W02617) but incorporates a KpnI site at
CC   codon 218, SstI site at codon 299, SstII at codon 333, BstBI at
CC   codon 338 and SalI immediately following the termination codon.
CC   These sites were incorporated by PCR to expedite construction of
CC   DNA sequences coding for p53 proteins bearing altered tetramerization
CC   domains or point mutations for use in cancer therapy.
XX
SQ   Sequence 1215 BP; 286 A; 382 C; 310 G; 237 T; 0 other;


alignment_scores:
            Quality:   55.00            Length:     11
              Ratio:  5.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x T32836    ..

 Align seg 1/1  to: T32836  from: 1  to: 1215

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||

     100 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 132

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V21730

seq_documentation_block:
ID   V21730 standard; cDNA; 1242 BP.
XX
AC   V21730;
XX
DT   17-AUG-1998  (first entry)
XX
DE   Human p53 cDNA.
XX
KW   Vector; vaccine; tumour; antigen; plasmid pITL-hHER/neu;
KW   human; p53; cancer; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   primer_bind     complement (360..386)
FT                   /*tag=  a
FT                   /note= "primer 1"
FT   primer_bind     985..1005
FT                   /*tag=  b
FT                   /note= "primer 2"
XX
PN   WO9806863-A1.
XX
PD   19-FEB-1998.
XX
PF   14-AUG-1997;   97WO-US14306.
XX
PR   14-AUG-1996;   96US-0023931.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Nelson EL,  Nelson PJ;
XX
DR   WPI; 1998-159552/14.
XX
PT   Humanised polynucleotide vectors - comprising human derived promoter
PT   and sequence acceptance site, used for the production of vaccines
XX
PS   Example 9; Page 90; 125pp; English.
XX
CC   This sequence comprises human p53 cDNA.  Primers (see V21728-29)
CC   from p53 cDNA can be used to target p53 sequences into novel
CC   humanised polynucleotide vectors such as plasmid pITL (see
CC   V21724).  These humanised vectors comprise a human-derived promoter
CC   (or mammalian homologue) which is functional in mammalian target
CC   tissue and cells and a sequence acceptance site which accepts cDNA
CC   products from RT-PCR cloning.  The vectors are non-replicating in
CC   mammalian cells but are capable of extended stable expression of
CC   the target sequence, generating an immune response in immunised
CC   individuals.  The vectors selectively elicit immune responses to
CC   the target sequences with little or no immune response to the other
CC   components of the vectors.  The vectors are particularly useful in
CC   accomodating monomorphic and polymorphic nucleic acid sequences
CC   encoding tumor antigens via PCR technology.
XX
SQ   Sequence 1242 BP; 282 A; 388 C; 320 G; 252 T; 0 other;


alignment_scores:
            Quality:   55.00            Length:     11
              Ratio:  5.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x V21730    ..

 Align seg 1/1  to: V21730  from: 1  to: 1242
```

```
      1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
        ||||||||||||||||||||||||||||||||
     73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X04533

seq_documentation_block:
ID   X04533 standard; DNA; 1303 BP.
XX
AC   X04533;
XX
DT   13-APR-1999  (first entry)
XX
DE   DNA encoding human p53 protein.
XX
KW   Ataxia telangiectasia; ATM protein; assay; interaction; kinase activity;
KW   p53; screening; ATR; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             122..1303
FT                   /*tag= a
FT                   /product= p53
XX
PN   GB2327498-A.
XX
PD   27-JAN-1999.
XX
PF   16-JUL-1998;   98GB-0015423.
XX
PR   16-JUL-1997;   97GB-0014971.
XX
PA   (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX
PI   Jackson SP,  Lakin ND,  Smith GCM;
XX
DR   WPI; 1999-073587/07.
DR   P-PSDB; W84270.
XX
PT   Assay method for compounds modulating the interaction of ATM and p53
PT   - useful for the treatment of e.g. cancer, immunosuppression and HIV
PT   infections and for the purification of the proteins ATM and ATR
XX
PS   Disclosure; Fig 7b; 124pp; English.
XX
CC   The present sequence encodes a human p53 protein. The protein is
CC   used in the assay of the invention. The specification describes an
CC   assay method for a compound able to modulate the interaction between
CC   ATM or a protein having an associated kinase activity and p53 or a
CC   protein having homologous phosphorylation sites. The assay comprises
CC   contacting a peptide fragment ATM with a relevant fragment of p53
CC   and a test compound, and determining the interaction or binding
CC   between the substances and the test compound. The assay method is
CC   useful for screening for compounds able to modulate the interaction
CC   between ATM and p53. The screened agents, peptide fragments and
CC   nucleic acids are useful for therapy involving modulating ATM action
CC   e.g. in the treatment of cancer, immunosuppression or HIV infections by
CC   modulating phosphorylation of p53 by ATM, and for purifying the proteins
CC   ATM and ATR.
XX
SQ   Sequence 1303 BP; 292 A; 403 C; 348 G; 260 T; 0 other;


alignment_scores:
           Quality:   55.00              Length:     11
             Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x X04533   ..

 Align seg 1/1  to: X04533  from: 1  to: 1303
```

```
      1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
        ||||||||||||||||||||||||||||||||
   194 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 226

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:D00088

seq_documentation_block:
ID   D00088 standard; DNA; 1307 BP.
XX
AC   D00088;
XX
DT   31-JUL-2000  (first entry)
XX
DE   Human tumour-associated antigen p53 DNA.
XX
KW   Human; tumour-associated antigen; p53 protein; DNA-binding domain;
KW   dermatological; immunosuppressive; antiinflammatory; autoimmune response;
KW   SLE; systemic lupus erythematosus; diagnosis; treatment; prevention; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             126..1307
FT                   /*tag= a
FT                   /product= "p53 protein"
XX
PN   WO200023082-A1.
XX
PD   27-APR-2000.
XX
PF   19-OCT-1999;   99WO-US24443.
XX
PR   19-OCT-1998;   98US-0104816.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Cohen IR,  Rotter V,  Erez-Alon N,  Herkel J;
XX
DR   WPI; 2000-339512/29.
DR   P-PSDB; Y70811.
XX
PT   Treatment of systemic lupus erythematosus by down-regulating the
PT   autoimmune response to the C-terminal DNA-binding domain of the p53
PT   protein by an active compound comprising of antibodies to p53 or
PT   fragments of p53 -
XX
PS   Disclosure; Page 79-81; 87pp; English.
XX
CC   The patent discloses a method for the treatment of systemic lupus
CC   erythematosus (SLE) by down-regulating the autoimmune response to the
CC   C-terminal DNA-binding domain of p53 protein by an active compound
CC   comprising C-terminal DNA-binding domain of p53, monoclonal
CC   antibodies (Ab1) specific to this domain, monoclonal antibodies (Ab2)
CC   specific to Ab1 and peptides based on the complementarity determining
CC   region of heavy and light chain of Ab1 and Ab2. The active compound is
CC   useful in the diagnosis, prevention and treatment of SLE in humans.
CC   The present sequence is a human tumour-associated antigen p53 DNA.
CC   Antibodies against the C-terminal DNA-binding domain of the p53
CC   protein can be raised and used for diagnosis and treatment of SLE.
XX
SQ   Sequence 1307 BP; 293 A; 404 C; 350 G; 260 T; 0 other;


alignment_scores:
           Quality:   55.00              Length:     11
             Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x D00088   ..

 Align seg 1/1  to: D00088  from: 1  to: 1307
```

```
  1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
    ||||||||||||||||||||||||||||||||
198 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 230
```

THIS PAGE BLANK (USPTO)

```
OM of: US-08-860-232-1 to: Issued_Patents_NA:*   out_format : pfs

Date: Dec 12, 2000  5:49 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US08860232/runat_11122000_154523_20748/app_query.fasta_1.67
 -DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=2000000000
 -USER=US08860232_@CGN1_1_43 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-08-860-232-1
Query length: 11
Database: Issued_Patents_NA:*
Database sequences: 262060
Database length: 75620496
Search time (sec): 57.120000



score_list:
Sequence           Strd Orig    ZScore   EScore Len  ! Documentation  ..
/cgn2_6/ptodata/1/ina/5D_COMB.seq:US-08-486-663A-18 +   55.00  174.89  0.0322    1181  !
/cgn2_6/ptodata/1/ina/5D_COMB.seq:US-08-247-904B-11 +   55.00  174.89  0.0322    1181  !
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-767-942A-22 +    55.00  174.89  0.0322    1181  !
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-184-009-215 +   55.00  174.88  0.0322    1182  !
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-484-956-92 +    55.00  174.88  0.0322    1182  !
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-484-956-93 +    55.00  174.88  0.0322    1182  !
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-484-956-94 +    55.00  174.88  0.0322    1182  !
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-757-653-92 +    55.00  174.88  0.0322    1182  !
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-757-653-93 +    55.00  174.88  0.0322    1182  !
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-757-653-94 +    55.00  174.88  0.0322    1182  !
/cgn2_6/ptodata/1/ina/5D_COMB.seq:US-08-458-356-215 +   55.00  174.88  0.0322    1182  !
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-347-792-20 +    55.00  174.62  0.0333    1215  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-431-357-20 +    55.00  174.62  0.0333    1215  !
/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US95-15353-20 +   55.00  174.62  0.0333    1215
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-047-041A-13 +   55.00  173.98  0.0361    1303  !
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-047-041A-12 +   55.00  173.95  0.0363    1307  !
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-07-912-011-1 +   55.00  173.88  0.0366    1317  !
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-347-792-1 +   55.00  173.88  0.0366    1317  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-431-357-1 +   55.00  173.88  0.0366    1317  !
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-697-221-1 +   55.00  173.88  0.0366    1317  !
/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US95-15353-1 +   55.00  173.88  0.0366    1317
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-184-009-103 +   55.00  172.78  0.0422    1484  !
/cgn2_6/ptodata/1/ina/5D_COMB.seq:US-08-458-356-103 +   55.00  172.78  0.0422    1484  !
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-184-009-99 +    55.00  172.60  0.0431    1512  !
/cgn2_6/ptodata/1/ina/5D_COMB.seq:US-08-458-356-99 +    55.00  172.60  0.0431    1512  !
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-126-587C-8 -    43.00  120.85  32.90     4032  !
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-690-495-8 -    40.00  145.17  1.45      91  !  Se
/cgn2_6/ptodata/1/ina/5D_COMB.seq:US-08-690-494-8 -    40.00  145.17  1.45      91  !  Se
/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US95-07671-8 -   40.00  145.17  1.45      91  !
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-690-495-9 -    40.00  140.87  2.52      145  !  S
/cgn2_6/ptodata/1/ina/5D_COMB.seq:US-08-690-494-9 -    40.00  140.87  2.52      145  !  S
/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US95-07671-9 -   40.00  140.87  2.52      145
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-094-071-1 +   38.50  112.51  95.93     1757  !  S
/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US92-00878-1 +   38.50  112.51  95.93     1757
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-781-891-79 -    38.00   74.67  1.2e+04   87350  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-404-445-2 -    37.00  112.63  94.38     972  !  S
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-404-445-1 -    37.00   96.01  795.41    5883  !
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-047-041A-16 +   36.00  144.04  1.68      22  !
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-047-041A-3 +   36.00  136.10  4.65      52  !  S
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-795-006A-23 +   36.00  136.10  4.65      52  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-472-239-1 +   36.00  110.87  118.29    800  !
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-722-184-7 -   36.00  106.44  208.83    1293  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-467-607-2 -   36.00  104.72  260.41    1558
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-469-362-2 -   36.00  104.72  260.41    1558
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-850-392-2 -   36.00  104.72  260.41    1558

seq_name: /cgn2_6/ptodata/1/ina/5D_COMB.seq:US-08-486-663A-18

seq_documentation_block:
; Sequence 18, Application US/08486663A
; Patent No. 5968761
;  GENERAL INFORMATION:
;    APPLICANT:  Rolfe, Mark
;    APPLICANT:  Eckstein, Jens W.
;    APPLICANT:  Draetta, Giulio
;    APPLICANT:  Guillaume Cottarel
;    TITLE OF INVENTION:  Ubiquitin Conjugating Enzymes
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII(text)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,663A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Vincent, Matthew P.
;      REGISTRATION NUMBER:  36,709
;      REFERENCE/DOCKET NUMBER:  MII-029CP2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1181 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..1182
US-08-486-663A-18


alignment_scores:
            Quality:   55.00              Length:      11
              Ratio:  5.000                Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-486-663A-18   ..

 Align seg 1/1  to: US-08-486-663A-18  from: 1  to: 1181

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /cgn2_6/ptodata/1/ina/5D_COMB.seq:US-08-247-904B-11
```

```
seq_documentation_block:
; Sequence 11, Application US/08247904B
; Patent No. 5981699
;  GENERAL INFORMATION:
;    APPLICANT:  Rolfe, Mark
;    APPLICANT:  Eckstein, Jens W.
;    APPLICANT:  Draetta, Giulio
;    TITLE OF INVENTION:  Human Ubiquitin Conjugating Enzyme
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley, Hoag & Eliot
;      STREET:  One Post Office Square
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII(text)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/247,904B
;      FILING DATE:  23-MAY-1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Vincent, Matthew P.
;      REGISTRATION NUMBER:  36,709
;      REFERENCE/DOCKET NUMBER:  MIV-029.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 832-1000
;      TELEFAX:  (617) 832-7000
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1181 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..1181
US-08-247-904B-11


alignment_scores:
            Quality:   55.00              Length:     11
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-247-904B-11    ..

 Align seg 1/1  to: US-08-247-904B-11  from: 1  to: 1181

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-767-942A-22

seq_documentation_block:
; Sequence 22, Application US/08767942A
; Patent No. 6068982
;  GENERAL INFORMATION:
;    APPLICANT:  Rolfe, Mark
;    APPLICANT:  Chiu, M. Isabel
;    APPLICANT:  Berlin, Vivian
;    APPLICANT:  Damagnez, Veronique
;    APPLICANT:  Draetta, Giulio
;    APPLICANT:  Guillaume, Cottarel
;    TITLE OF INVENTION:  UBIQUITIN CONJUGATING ENZYMES
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY, HOAG & ELIOT LLP
;      STREET:  One Post Office Square
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109-2170
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/767,942A
;      FILING DATE:  17-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Vincent, Matthew P.
;      REGISTRATION NUMBER:  36,709
;      REFERENCE/DOCKET NUMBER:  MIV-029.04
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-832-1000
;      TELEFAX:  617-832-7000
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1181 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..1179
US-08-767-942A-22


alignment_scores:
            Quality:   55.00              Length:     11
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-767-942A-22    ..

 Align seg 1/1  to: US-08-767-942A-22  from: 1  to: 1181

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-184-009-215

seq_documentation_block:
; Sequence 215, Application US/08184009
; Patent No. 5833975
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Tartaglia, James
;    APPLICANT:  Cox, William I.
;    TITLE OF INVENTION:  RECOMBINANT VIRUS IMMUNOTHERAPY
;    NUMBER OF SEQUENCES:  217
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/184,009
;      FILING DATE:  19-JAN-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2530
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;      TELEX:  425066CURTMS
;  INFORMATION FOR SEQ ID NO:  215:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1182 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-184-009-215


alignment_scores:
            Quality:   55.00              Length:     11
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-184-009-215    ..

 Align seg 1/1  to: US-08-184-009-215  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-484-956-92

seq_documentation_block:
; Sequence 92, Application US/08484956
; Patent No. 5843654
;  GENERAL INFORMATION:
;    APPLICANT:  DAHLBERG, JAMES E.
;    APPLICANT:  LYAMICHEV, VICTOR I.
;    APPLICANT:  BROW, MARY ANN D.
;    APPLICANT:  OLDENBURG, MARY C.
;    APPLICANT:  HEISLER, LAURA
;    TITLE OF INVENTION:  DETECTION OF p53 MUTATIONS
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HAVERSTOCK, MEDLEN & CARROLL
;      STREET:  220 MONTGOMERY STREET, SUITE 2200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,956
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/402,601
;      FILING DATE:  09-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/337,164
;      FILING DATE:  09-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/254,359
;      FILING DATE:  06-JUN-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/073,384
;      FILING DATE:  04-JUN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/986,330
;      FILING DATE:  07-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CARRPOLL ], PETER G.
;      REGISTRATION NUMBER:  32,837
;      REFERENCE/DOCKET NUMBER:  FORS-01801
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  92:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1182 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-484-956-92


alignment_scores:
            Quality:   55.00              Length:     11
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-484-956-92    ..

 Align seg 1/1  to: US-08-484-956-92  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-484-956-93

seq_documentation_block:
; Sequence 93, Application US/08484956
; Patent No. 5843654
;  GENERAL INFORMATION:
;    APPLICANT:  DAHLBERG, JAMES E.
;    APPLICANT:  LYAMICHEV, VICTOR I.
;    APPLICANT:  BROW, MARY ANN D.
;    APPLICANT:  OLDENBURG, MARY C.
;    APPLICANT:  HEISLER, LAURA
;    TITLE OF INVENTION:  DETECTION OF p53 MUTATIONS
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HAVERSTOCK, MEDLEN & CARROLL
;      STREET:  220 MONTGOMERY STREET, SUITE 2200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,956
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/402,601
;      FILING DATE:  09-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/337,164
;      FILING DATE:  09-NOV-1994
;    PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US 08/254,359
;       FILING DATE:  06-JUN-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/073,384
;       FILING DATE:  04-JUN-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/986,330
;       FILING DATE:  07-DEC-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  CARRPOLL ], PETER G.
;       REGISTRATION NUMBER:  32,837
;       REFERENCE/DOCKET NUMBER:  FORS-01801
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 705-8410
;       TELEFAX:  (415) 397-8338
;   INFORMATION FOR SEQ ID NO:  93:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1182 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-484-956-93


alignment_scores:
            Quality:   55.00              Length:     11
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-484-956-93    ..

 Align seg 1/1  to: US-08-484-956-93  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-484-956-94

seq_documentation_block:
; Sequence 94, Application US/08484956
; Patent No. 5843654
;  GENERAL INFORMATION:
;     APPLICANT:  DAHLBERG, JAMES E.
;     APPLICANT:  LYAMICHEV, VICTOR I.
;     APPLICANT:  BROW, MARY ANN D.
;     APPLICANT:  OLDENBURG, MARY C.
;     APPLICANT:  HEISLER, LAURA
;     TITLE OF INVENTION:  DETECTION OF p53 MUTATIONS
;     NUMBER OF SEQUENCES:  114
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  HAVERSTOCK, MEDLEN & CARROLL
;       STREET:  220 MONTGOMERY STREET, SUITE 2200
;       CITY:  SAN FRANCISCO
;       STATE:  CALIFORNIA
;       COUNTRY:  UNITED STATES OF AMERICA
;       ZIP:  94104
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/484,956
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/402,601
;       FILING DATE:  09-MAR-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/337,164
;       FILING DATE:  09-NOV-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/254,359
;       FILING DATE:  06-JUN-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/073,384
;       FILING DATE:  04-JUN-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/986,330
;       FILING DATE:  07-DEC-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  CARRPOLL ], PETER G.
;       REGISTRATION NUMBER:  32,837
;       REFERENCE/DOCKET NUMBER:  FORS-01801
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 705-8410
;       TELEFAX:  (415) 397-8338
;   INFORMATION FOR SEQ ID NO:  94:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1182 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-484-956-94


alignment_scores:
            Quality:   55.00              Length:     11
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-484-956-94    ..

 Align seg 1/1  to: US-08-484-956-94  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-757-653-92

seq_documentation_block:
; Sequence 92, Application US/08757653
; Patent No. 5843669
;  GENERAL INFORMATION:
;     APPLICANT:  Kaiser, Michael W.
;     APPLICANT:  Lyamichev, Victor I.
;     APPLICANT:  Lyamichev, Natasha
;     TITLE OF INVENTION:  Cleavage Of Nucleic Acid Using
;     TITLE OF INVENTION:  Thermostable FEN-1 Endonucleases
;     NUMBER OF SEQUENCES:  190
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Medlen & Carroll, LLP
;       STREET:  220 Montgomery Street, Suite 2200
;       CITY:  San Francisco
;       STATE:  California
;       COUNTRY:  United States Of America
;       ZIP:  94104
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/757,653
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Ingolia, Diane E.
;       REGISTRATION NUMBER:  40,027
;       REFERENCE/DOCKET NUMBER:  FORS-02565
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  92:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1182 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-757-653-92


alignment_scores:
            Quality:   55.00              Length:     11
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-757-653-92    ..

 Align seg 1/1  to: US-08-757-653-92  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-757-653-93

seq_documentation_block:
; Sequence 93, Application US/08757653
; Patent No. 5843669
;  GENERAL INFORMATION:
;    APPLICANT:  Kaiser, Michael W.
;    APPLICANT:  Lyamichev, Victor I.
;    APPLICANT:  Lyamichev, Natasha
;    TITLE OF INVENTION:  Cleavage Of Nucleic Acid Using
;    TITLE OF INVENTION:  Thermostable FEN-1 Endonucleases
;    NUMBER OF SEQUENCES:  190
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States Of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/757,653
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ingolia, Diane E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  FORS-02565
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  93:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1182 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-757-653-93


alignment_scores:
            Quality:   55.00              Length:     11
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-757-653-93    ..

 Align seg 1/1  to: US-08-757-653-93  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-757-653-94

seq_documentation_block:
; Sequence 94, Application US/08757653
; Patent No. 5843669
;  GENERAL INFORMATION:
;    APPLICANT:  Kaiser, Michael W.
;    APPLICANT:  Lyamichev, Victor I.
;    APPLICANT:  Lyamichev, Natasha
;    TITLE OF INVENTION:  Cleavage Of Nucleic Acid Using
;    TITLE OF INVENTION:  Thermostable FEN-1 Endonucleases
;    NUMBER OF SEQUENCES:  190
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States Of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/757,653
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ingolia, Diane E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  FORS-02565
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  94:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1182 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-757-653-94


alignment_scores:
            Quality:   55.00              Length:     11
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-757-653-94    ..

 Align seg 1/1  to: US-08-757-653-94  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /cgn2_6/ptodata/1/ina/5D_COMB.seq:US-08-458-356-215
```

```
seq_documentation_block:
; Sequence 215, Application US/08458356
; Patent No. 5942235
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Tartaglia, James
;    APPLICANT:  Cox, William I.
;    TITLE OF INVENTION:  RECOMBINANT VIRUS IMMUNOTHERAPY
;    NUMBER OF SEQUENCES:  217
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,356
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/184,009
;      FILING DATE:  19-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2530
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;      TELEX:  425066CURTMS
;  INFORMATION FOR SEQ ID NO:  215:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1182 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-458-356-215


alignment_scores:
            Quality:   55.00              Length:     11
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-458-356-215    ..

 Align seg 1/1  to: US-08-458-356-215  from: 1  to: 1182

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||||||||||||||||||||||||||
      73 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 105

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-347-792-20

seq_documentation_block:
; Sequence 20, Application US/08347792
; Patent No. 5573925
;  GENERAL INFORMATION:
;    APPLICANT:  Halazonetis, Thanos D.
;    TITLE OF INVENTION:  p53 Proteins With Altered
;    TITLE OF INVENTION:  Tetramerization Domains
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr., PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/347,792
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST58USA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9206
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1215 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-347-792-20


alignment_scores:
            Quality:   55.00              Length:     11
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-347-792-20    ..

 Align seg 1/1  to: US-08-347-792-20  from: 1  to: 1215

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||||||||||||||||||||||||||
     100 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 132

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-431-357-20

seq_documentation_block:
; Sequence 20, Application US/08431357
; Patent No. 5721340
;  GENERAL INFORMATION:
;    APPLICANT:  Halazonetis, Thanos D.
;    TITLE OF INVENTION:  p53 Proteins With Altered
;    TITLE OF INVENTION:  Tetramerization Domains
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr., PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/431,357
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/347,792
```

```
;       FILING DATE:  28-NOV-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Bak, Mary E.
;       REGISTRATION NUMBER:  31,215
;       REFERENCE/DOCKET NUMBER:  WST58USA
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  215-540-9206
;       TELEFAX:  215-540-5818
;   INFORMATION FOR SEQ ID NO:  20:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1215 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-431-357-20


alignment_scores:
            Quality:   55.00              Length:     11
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-431-357-20   ..

 Align seg 1/1  to: US-08-431-357-20  from: 1  to: 1215

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||||||||||||||||||||||||||
     100 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 132

seq_name: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US95-15353-20

seq_documentation_block:
; Sequence 20, Application PC/TUS9515353
;  GENERAL INFORMATION:
;     APPLICANT:  The Wistar Institute of Anatomy
;     APPLICANT:  and Biology
;     APPLICANT:  Halazonetis, Thanos D.
;     TITLE OF INVENTION:  p53 Proteins With Altered
;     TITLE OF INVENTION:  Tetramerization Domains
;     NUMBER OF SEQUENCES:  46
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Howson and Howson
;       STREET:  Spring House Corporate Cntr., PO Box 457
;       CITY:  Spring House
;       STATE:  Pennsylvania
;       COUNTRY:  USA
;       ZIP:  19477
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/15353
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/347,792
;       FILING DATE:  28-NOV-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/431,357
;       FILING DATE:  28-APR-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/456,623
;       FILING DATE:  01-JUN-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Bak, Mary E.
;       REGISTRATION NUMBER:  31,215
;       REFERENCE/DOCKET NUMBER:  WST58CPCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  215-540-9206
;       TELEFAX:  215-540-5818
;   INFORMATION FOR SEQ ID NO:  20:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1215 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
PCT-US95-15353-20


alignment_scores:
            Quality:   55.00              Length:     11
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x PCT-US95-15353-20   ..

 Align seg 1/1  to: PCT-US95-15353-20  from: 1  to: 1215

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||||||||||||||||||||||||||
     100 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 132

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-047-041A-13

seq_documentation_block:
; Sequence 13, Application US/08047041A
; Patent No. 5527676
;  GENERAL INFORMATION:
;     APPLICANT:  Vogelstein, Bert
;     APPLICANT:  Baker, Suzanne J.
;     APPLICANT:  Fearon, Eric R.
;     APPLICANT:  Nigro, Janice M.
;     TITLE OF INVENTION:  Detection of Loss of the Wild-Type p53
;     TITLE OF INVENTION:  Gene
;     NUMBER OF SEQUENCES:  28
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Banner & Allegretti, Ltd.
;       STREET:  1001 G Street, N.W.
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  USA
;       ZIP:  20001.4597
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/047,041A
;       FILING DATE:  22-MAR-1993
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/928,661
;       FILING DATE:  17-AUG-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/446,584
;       FILING DATE:  06-DEC-1989
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/330,566
;       FILING DATE:  29-MAR-1989
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kagan, Sarah A.
;       REGISTRATION NUMBER:  32,141
;       REFERENCE/DOCKET NUMBER:  01107.42917
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-508-9100
;       TELEFAX:  202-508-9299
;   INFORMATION FOR SEQ ID NO:  13:
;     SEQUENCE CHARACTERISTICS:
```

```
;       LENGTH:  1303 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapiens
;     POSITION IN GENOME:
;       MAP POSITION:  17p13.1
;     PUBLICATION INFORMATION:
;       AUTHORS:  Harris, N.
;       JOURNAL:  Mol. Cell. Biol
;       VOLUME:  6
;       ISSUE:  12
;       PAGES:  4650-4656
;       DATE:  1986
US-08-047-041A-13


alignment_scores:
            Quality:   55.00              Length:     11
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-047-041A-13   ..

 Align seg 1/1  to: US-08-047-041A-13  from: 1  to: 1303

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||||||||||||||||||||||
     194 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 226
```

```
OM of: US-08-860-232-1 to: EST:*   out_format : pfs

Date: Dec 12, 2000  4:17 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US08860232/runat_11122000_154522_20724/app_query.fasta_1.67
 -DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US08860232_@CGN1_1_1780 -NCPU=6
 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-08-860-232-1
Query length: 11
Database: EST:*
Database sequences: 7189864
Database length: -1203564053
Search time (sec): 437.680000



score_list:
Sequence              Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_est22:AW407968      +    55.00  182.86  0.3210    401  ! AW407968 UI-HF-BM0-adl-a-03-0-U
gb_est2:AA171861       +    55.00  182.53  0.3348    416  ! AA171861 zp22c06.r1 Stratagene
gb_gss13:AQ887186      +    46.00  146.06  35.99     694  ! AQ887186 HS_5551_B2_E06_T7A RPC
gb_est7:AA898573       +    45.00  143.83  47.93     599  ! AA898573 NCP2B7T3 Perithecial N
gb_est13:AI894873      +    44.00  141.03  68.66     551  ! AI894873 EST264316 tomato callu
gb_est19:AW030386      +    44.00  140.68  71.80     573  ! AW030386 EST273641 tomato callu
gb_est19:AW030009      +    44.00  140.66  71.94     574  ! AW030009 EST273264 tomato callu
gb_gss17:AZ185470      -    44.00  138.45  95.57     736  ! AZ185470 SP_1005_B1_A06_SP6E St
gb_est38:R44262        +    43.00  143.02  53.21     296  ! R44262 yg35e08.s1 Soares infant
gb_est5:AA687270       +    42.00  136.33  125.38    421  ! AA687270 nv59f06.s1 NCI_CGAP_GC
gb_gss1:AQ056407       -    42.00  135.62  137.36    456  ! AQ056407 CIT-HSP-2342L24.TF CIT
gb_est33:BE138282      -    42.00  135.41  141.15    467  ! BE138282 ug50g06.y1 Barstead bo
gb_est11:AI510593      -    42.00  134.80  152.59    500  ! AI510593 mq40a07.y1 Barstead MP
gb_est1:AA139619       -    42.00  134.38  160.99    524  ! AA139619 mq40a07.r1 Barstead MP
gb_est25:AW975473      +    42.00  132.59  202.66    641  ! AW975473 EST387582 MAGE reseque
gb_est36:BE540270      -    42.00  131.92  220.82    691  ! BE540270 601059844F1 NIH_MGC_10
gb_est35:BE366393      +    42.00  131.65  228.50    712  ! BE366393 PI1_32_G07.g1_A002 Pat
gb_gss22:CNS024JQ      -    42.00  128.35  349.14   1032  ! AL180863 Tetraodon nigroviridi
gb_est24:AW799589      +    41.00  144.99  41.33     107  ! AW799589 PM2-UM0053-160300-004-
gb_gss19:B15996        -    41.00  134.51  158.43    347  ! B15996 346K8.TV CIT978SKA1 Homo
gb_est20:AW147745      -    41.00  133.38  183.17    394  ! AW147745 da16f12.y1 normalized
gb_gss15:AZ045560      -    41.00  132.52  204.56    434  ! AZ045560 T234221b shotgun sub-1
gb_gss15:AZ045581      -    41.00  132.07  216.45    456  ! AZ045581 T234248b shotgun sub-1
gb_est20:AW147256      +    41.00  131.96  219.70    462  ! AW147256 da16f12.x1 normalized
gb_gss15:AZ074230      -    41.00  131.94  220.25    463  ! AZ074230 RPCI-23-395J12.TJ RPCI
gb_est14:AL120309      +    41.00  131.45  234.43    489  ! AL120309 DKFZp761N247_r1 761 (s
gb_gss15:AZ067525      -    41.00  129.77  291.06    591  ! AZ067525 RPCI-23-432M5.TV RPCI-
gb_gss18:AZ256714      -    41.00  128.52  341.64    680  ! AZ256714 RPCI-23-133P6.TJ RPCI-
gb_est35:BE456079      +    41.00  126.19  460.43    883  ! BE456079 HVSMEg0019J24f Hordeum
gb_est12:AI656536      +    40.00  133.46  181.12    262  ! AI656536 tt46d05.x1 NCI_CGAP_GC
gb_est26:BB025750      +    40.00  132.97  193.01    277  ! BB025750 BB025750 RIKEN full-le
gb_est19:AW090465      +    40.00  132.75  198.59    284  ! AW090465 xc83e02.x1 NCI_CGAP_Br
gb_est36:C01227        -    40.00  132.71  199.39    285  ! C01227 HUMGS0007946 Human adult
gb_est28:BB238903      +    40.00  131.68  227.60    320  ! BB238903 BB238903 RIKEN full-le
gb_est1:AA064166       +    40.00  131.60  230.04    323  ! AA064166 mj62h06.r1 Soares mous
gb_est23:AW662994      -    40.00  131.25  240.64    336  ! AW662994 hh83h05.y1 NCI_CGAP_GU
gb_est36:C70820        +    40.00  130.63  260.37    360  ! C70820 C70820 Yuji Kohara unpub
gb_gss8:AQ567801       +    40.00  130.61  261.20    361  ! AQ567801 HS_2116_A2_E04_MR CIT
gb_est37:H72801        -    40.00  130.54  263.68    364  ! H72801 yu07d11.r1 Soares fetal
gb_est8:AI089900       +    40.00  130.37  269.48    371  ! AI089900 qa16a04.x1 NCI_CGAP_Br
gb_est10:AI439133      +    40.00  130.32  271.14    373  ! AI439133 ti87e02.x1 NCI_CGAP_Ki
gb_est13:AI911403      +    40.00  130.27  272.80    375  ! AI911403 wd16g01.x1 Soares_N
gb_est10:AI334605      +    40.00  130.15  276.96    380  ! AI334605 tb20c06.x1 NCI_CGAP
gb_est11:AI479512      +    40.00  130.04  281.12    385  ! AI479512 tm46e02.x1 NCI_CGAP
gb_est9:AI306677       +    40.00  129.92  285.30    390  ! AI306677 qw25a05.x1 NCI_CGAP

seq_name: gb_est22:AW407968

seq_documentation_block:
LOCUS       AW407968      401 bp    mRNA            EST       16-FEB-2000
DEFINITION  UI-HF-BM0-adl-a-03-0-UI.r2 NIH_MGC_38 Homo sapiens cDNA clone
            IMAGE:3061901 5', mRNA sequence.
ACCESSION   AW407968
VERSION     AW407968.1  GI:6927025
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 401)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Eco RI site shown at the beginning of the sequence.
             Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: M.B. Soares Lab
             cDNA Library Arrayed by: M.B. Soares Lab
             DNA Sequencing by: M.B. Soares Lab
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1. .401
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3061901"
                     /clone_lib="NIH_MGC_38"
                     /tissue_type="lymph"
                     /cell_type="germinal center B cells"
                     /cell_line="MGC85"
                     /lab_host="DH10B (LTI)"
                     /note="Vector: pT7T3-Pac; Site_1: NotI; Site_2: Eco RI;
                     Constructed from size fractionated cytoplasmic mRNA
                     (2.5-3.5kb). Directionally cloned. Cells provided by Louis
                     M. Staudt, Ph.D. Library preparation by Maria de Fatima
                     Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT       83 a    140 c     96 g     82 t
ORIGIN


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x AW407968   ..

 Align seg 1/1  to: AW407968  from: 1  to: 401

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||||||||||||||||||||||||||||
     136 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 168

seq_name: gb_est2:AA171861

seq_documentation_block:
LOCUS       AA171861      416 bp    mRNA            EST       23-DEC-1996
DEFINITION  zp22c06.r1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA
            clone IMAGE:610186 5' similar to gb:X54156_rna1 CELLULAR TUMOR
```

```
            ANTIGEN P53 (HUMAN);, mRNA sequence.
ACCESSION   AA171861
VERSION     AA171861.1  GI:1750919
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 416)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28M13 rev2 from Amersham.
FEATURES             Location/Qualifiers
     source          1. .416
                     /organism="Homo sapiens"
                     /db_xref="GDB:4625445"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:610186"
                     /clone_lib="Stratagene neuroepithelium (#937231)"
                     /dev_stage="Ntera-2/RA neuroepithelial cells"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. NT2
                     cells (Ntera-2/c1.D1) induced with Retinoic Acid for 24
                     hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       80 a    138 c    110 g     86 t      2 others
ORIGIN


alignment_scores:
            Quality:   55.00             Length:     11
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x AA171861   ..

 Align seg 1/1  to: AA171861  from: 1  to: 416

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||||||||||||||||||||||||
     198 CTACTTCCTGAAAACAACGTTCTGTCCCCCTTG 230

seq_name: gb_gss13:AQ887186

seq_documentation_block:
LOCUS       AQ887186      694 bp    DNA             GSS       10-NOV-1999
DEFINITION  HS_5551_B2_E06_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=9319 Col=12 Row=J, DNA sequence.
ACCESSION   AQ887186
VERSION     AQ887186.1  GI:6343376
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 694)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 9319  row: J  column: 12
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 694.
FEATURES             Location/Qualifiers
     source          1. .694
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=9319 Col=12 Row=J"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT      177 a    164 c    137 g    189 t     27 others
ORIGIN


alignment_scores:
            Quality:   46.00             Length:     10
              Ratio:  4.600               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 80.000

alignment_block:
 US-08-860-232-1 x AQ887186   ..

 Align seg 1/1  to: AQ887186  from: 1  to: 694

       2 LeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||||::::::|||||||||
     625 CTCCCGGAGAACAACATCTTCTCGCCACTT 654

seq_name: gb_est7:AA898573

seq_documentation_block:
LOCUS       AA898573      599 bp    mRNA            EST       12-APR-1998
DEFINITION  NCP2B7T3 Perithecial Neurospora crassa cDNA clone NP2B7 5' end,
            mRNA sequence.
ACCESSION   AA898573
VERSION     AA898573.1  GI:3045006
KEYWORDS    EST.
SOURCE      Neurospora crassa.
  ORGANISM  Neurospora crassa
            Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae;
            Neurospora.
REFERENCE   1  (bases 1 to 599)
  AUTHORS   Nelson,M.A., Kang,S., Braun,E.L., Crawford,M.E., Dolan,P.L.,
            Leonard,P.M., Mitchell,J., Armijo,A.M., Bean,L., Blueyes,E.,
            Cushing,T., Errett,A., Fleharty,M., Gorman,M., Judson,K., Miller,R.
            , Ortega,J., Pavlova,I., Perea,J., Todisco,S., Trujillo,R.,
```

```
            Valentine,J., Wells,A., Werner-Washburne,M., Yazzie,S. and Natvig
            ,D.O.
  TITLE     Expressed sequences from conidial, mycelial, and sexual stages of
            Neurospora crassa
  JOURNAL   Fungal Genet. Biol. 21, 348-363 (1997)
  MEDLINE   97435549
COMMENT     Contact: Natvig,D.O./Nelson,M.A.
            Department of Biology
            University of New Mexico
            Castetter Hall, Albuquerque, NM 87131, USA
            Tel: 505 277 3411
            Fax: 505 277 0304
            Email: ngp@biology.unm.edu
            Deposited in GSDB at the National Center for Genome Resources with
            accession GSDB:S:1147504
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .599
                     /organism="Neurospora crassa"
                     /strain="fl a"
                     /db_xref="taxon:5141"
                     /clone="NP2B7"
                     /clone_lib="Perithecial"
                     /sex="Mating type a (fluffy), fertilized"
                     /tissue_type="Perithecia (fruiting bodies)"
                     /dev_stage="perithecia"
                     /note="mRNA isolated from 5 day old perithecia (fruiting
                     bodies) of the fluffy strain fl a (Mating type a),
                     fertilized with conidia from 74-OR23-IV A (Mating type A).
                     cDNA directionally cloned into pBluescript SK(-) using the
                     Uni-ZAP XR vector system (Stratagene, La Jolla, CA)."
BASE COUNT      129 a    219 c    107 g    144 t
ORIGIN


alignment_scores:
            Quality:   45.00            Length:     10
              Ratio:  4.500              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 90.000

alignment_block:
 US-08-860-232-1 x AA898573   ..

 Align seg 1/1  to: AA898573  from: 1  to: 599

       2 LeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||:::||||||||||||||||||
     383 CTGCCAGAGACCAACGTCTTGTCCCCTCTT 412

seq_name: gb_est13:AI894873

seq_documentation_block:
LOCUS       AI894873      551 bp    mRNA            EST       27-JUL-1999
DEFINITION  EST264316 tomato callus, TAMU Lycopersicon esculentum cDNA clone
            cLEC6K7, mRNA sequence.
ACCESSION   AI894873
VERSION     AI894873.1  GI:5600775
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
            I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 551)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Ronning
            ,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato callus tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: David Frisch
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 4366
            Fax: 864 656 4293
            Email: dfrisch@CLEMSON.EDU
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .551
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEC6K7"
                     /clone_lib="tomato callus, TAMU"
                     /tissue_type="callus"
                     /dev_stage="25-40 days old"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; supplier: Giovannoni laboratory; cLEC - Cotyledons
                     of seedlings 7-10 days post-germination were excised, cut
                     at both ends and placed on MS medium with no selection.
                     Mixed callus was harvested at 25 and 40 days and included
                     undifferentiated masses. Tomato Callus EST Library"
BASE COUNT      180 a     87 c    118 g    166 t
ORIGIN


alignment_scores:
            Quality:   44.00            Length:     11
              Ratio:  4.400              Gaps:      0
 Percent Similarity:  90.909   Percent Identity: 72.727

alignment_block:
 US-08-860-232-1 x AI894873   ..

 Align seg 1/1  to: AI894873  from: 1  to: 551

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||:::||||||:::|||   ||||||
      75 TTACTCCCTAGAAATAATATTCTTCGTCCATTA 107

seq_name: gb_est19:AW030386

seq_documentation_block:
LOCUS       AW030386      573 bp    mRNA            EST       15-SEP-1999
DEFINITION  EST273641 tomato callus, TAMU Lycopersicon esculentum cDNA clone
            cLEC20I12, mRNA sequence.
ACCESSION   AW030386
VERSION     AW030386.1  GI:5889142
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
            I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Ronning
            ,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato callus tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: David Frisch
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 4366
            Fax: 864 656 4293
            Email: dfrisch@CLEMSON.EDU
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEC20I12"
                     /clone_lib="tomato callus, TAMU"
```

```
                     /tissue_type="callus"
                     /dev_stage="25-40 days old"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; supplier: Giovannoni laboratory; cLEC - Cotyledons
                     of seedlings 7-10 days post-germination were excised, cut
                     at both ends and placed on MS medium with no selection.
                     Mixed callus was harvested at 25 and 40 days and included
                     undifferentiated masses. Tomato Callus EST Library"
BASE COUNT      191 a    106 c    106 g    170 t
ORIGIN


alignment_scores:
            Quality:   44.00             Length:     11
              Ratio:  4.400               Gaps:      0
 Percent Similarity:  90.909   Percent Identity: 72.727

alignment_block:
 US-08-860-232-1 x AW030386   ..

 Align seg 1/1  to: AW030386  from: 1  to: 573

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||:::||||||:::|||   ||||||
     208 TTACTCCCTAGAAATAATATTCTTCGTCCATTA 240

seq_name: gb_est19:AW030009

seq_documentation_block:
LOCUS       AW030009      574 bp    mRNA            EST       15-SEP-1999
DEFINITION  EST273264 tomato callus, TAMU Lycopersicon esculentum cDNA clone
            cLEC11J16, mRNA sequence.
ACCESSION   AW030009
VERSION     AW030009.1  GI:5888765
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
            I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 574)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Ronning
            ,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato callus tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: David Frisch
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 4366
            Fax: 864 656 4293
            Email: dfrisch@CLEMSON.EDU
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .574
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEC11J16"
                     /clone_lib="tomato callus, TAMU"
                     /tissue_type="callus"
                     /dev_stage="25-40 days old"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; supplier: Giovannoni laboratory; cLEC - Cotyledons
                     of seedlings 7-10 days post-germination were excised, cut
                     at both ends and placed on MS medium with no selection.
                     Mixed callus was harvested at 25 and 40 days and included
                     undifferentiated masses. Tomato Callus EST Library"
BASE COUNT      188 a     97 c    118 g    170 t      1 others
ORIGIN


alignment_scores:
            Quality:   44.00             Length:     11
              Ratio:  4.400               Gaps:      0
 Percent Similarity:  90.909   Percent Identity: 72.727

alignment_block:
 US-08-860-232-1 x AW030009   ..

 Align seg 1/1  to: AW030009  from: 1  to: 574

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||:::||||||:::|||   ||||||
     136 TTACTCCCTAGAAATAATATTCTTCGTCCATTA 168

seq_name: gb_gss17:AZ185470

seq_documentation_block:
LOCUS       AZ185470      736 bp    DNA             GSS       08-JUN-2000
DEFINITION  SP_1005_B1_A06_SP6E Strongylocentrotus purpuratus, purple sea
            urchin, sperm genomic BAC library Strongylocentrotus purpuratus
            genomic clone Plate=1005 Col=11 Row=B, DNA sequence.
ACCESSION   AZ185470
VERSION     AZ185470.1  GI:8357947
KEYWORDS    GSS.
SOURCE      Strongylocentrotus purpuratus.
  ORGANISM  Strongylocentrotus purpuratus
            Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
            Echinoidea; Euechinoidea; Echinacea; Echinoida;
            Strongylocentrotidae; Strongylocentrotus.
REFERENCE   1  (bases 1 to 736)
  AUTHORS   Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,
            Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray
            ,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J, Davidson,E.H. and
            Hood,L.
  TITLE     A Sea Urchin Genome Project: Sequence Scan, Virtual Map, and
            Additional Resources
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cameron, RA, Davidson, EH, Hood, L
            Division of Biology 156-29
            California Institute of Technology
            Pasadena California 91125, USA
            Tel: (626) 395-8421
            Fax: (626) 793-3047
            Email: acameron@caltech.edu
            Plate: 1005  row: B  column: 11
            Seq primer: SP6
            Class: BAC ends
            High quality sequence stop: 736.
FEATURES             Location/Qualifiers
     source          1. .736
                     /organism="Strongylocentrotus purpuratus"
                     /db_xref="taxon:7668"
                     /clone="Plate=1005 Col=11 Row=B"
                     /clone_lib="Strongylocentrotus purpuratus, purple sea
                     urchin, sperm genomic BAC library"
                     /note="Organ: sperm; Vector: BACe3.6; BAC Clones in E-Coli
                     DH10B"
BASE COUNT      224 a    129 c    131 g    250 t      2 others
ORIGIN


alignment_scores:
            Quality:   44.00             Length:     11
              Ratio:  4.400               Gaps:      0
 Percent Similarity:  90.909   Percent Identity: 63.636

alignment_block:
 US-08-860-232-1 x AZ185470/rev   ..

 Align seg 1/1  to reverse of: AZ185470  from: 1  to: 736
```

```
      1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
        ::::::|||   ||||||:::||||||||||||
    665 ATTATTCCAACAAATAATATATTATCCCCTCTT 633

seq_name: gb_est38:R44262

seq_documentation_block:
LOCUS       R44262       296 bp    mRNA            EST       22-MAY-1995
DEFINITION  yg35e08.s1 Soares infant brain 1NIB Homo sapiens cDNA clone
            IMAGE:34286 3', mRNA sequence.
ACCESSION   R44262
VERSION     R44262.1  GI:820620
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 296)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     On May 9, 1995 this sequence version replaced gi:802986.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 813
            High quality sequence stops: 146 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 813   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 146.
FEATURES             Location/Qualifiers
     source          1. .296
                     /organism="Homo sapiens"
                     /db_xref="GDB:406633"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:34286"
                     /clone_lib="Soares infant brain 1NIB"
                     /sex="female"
                     /dev_stage="73 days post natal"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
                     I; Site_2: Hind III; 1st strand cDNA was primed with a Not
                     I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Hind III adaptors
                     (Pharmacia), digested with Not I and directionally cloned
                     into the Not I and Hind III sites of the Lafmid BA vector.
                     Library went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       92 a     47 c     34 g    116 t      7 others
ORIGIN


alignment_scores:
            Quality:   43.00            Length:     10
              Ratio:  4.778              Gaps:      0
 Percent Similarity:  90.000   Percent Identity: 80.000

alignment_block:
 US-08-860-232-1 x R44262   ..

 Align seg 1/1  to: R44262  from: 1  to: 296

      1 LeuLeuProGluAsnAsnValLeuSerPro 10
        ||||||||||||||||||   :::|||||||
    146 TTGTTGCCGGAAAACAACAAAATATCCCCA 175

seq_name: gb_est5:AA687270

seq_documentation_block:
LOCUS       AA687270      421 bp    mRNA            EST       24-DEC-1997
DEFINITION  nv59f06.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1234115 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AA687270
VERSION     AA687270.1  GI:2675461
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 421)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            unknown library type
            Insert Length: 677   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .421
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1234115"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      145 a     82 c     77 g    117 t
ORIGIN


alignment_scores:
            Quality:   42.00            Length:     11
              Ratio:  4.200              Gaps:      0
 Percent Similarity:  90.909   Percent Identity: 72.727

alignment_block:
 US-08-860-232-1 x AA687270   ..

 Align seg 1/1  to: AA687270  from: 1  to: 421

      1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
        |||||||||:::|||||||||:::   ||||||
    367 CTCCTCCCCCATAATAATGTCATGCTGCCTCTA 399

seq_name: gb_gss1:AQ056407

seq_documentation_block:
LOCUS       AQ056407      456 bp    DNA             GSS       30-JUL-1998
DEFINITION  CIT-HSP-2342L24.TF CIT-HSP Homo sapiens genomic clone 2342L24, DNA
            sequence.
ACCESSION   AQ056407
VERSION     AQ056407.1  GI:3353013
```

```
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2342L24"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      130 a    102 c    125 g     99 t
ORIGIN


alignment_scores:
            Quality:   42.00             Length:     11
              Ratio:  4.200               Gaps:      0
 Percent Similarity:  90.909   Percent Identity: 72.727

alignment_block:
 US-08-860-232-1 x AQ056407/rev   ..

 Align seg 1/1  to reverse of: AQ056407  from: 1  to: 456

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||   :::|||:::||||||||||||
     195 CTGCTCCCAGCCAGGAACATTTTAAGTCCTCTA 163

seq_name: gb_est33:BE138282

seq_documentation_block:
LOCUS       BE138282      467 bp    mRNA            EST       21-JUN-2000
DEFINITION  ug50g06.y1 Barstead bowel MPLRB9 Mus musculus cDNA clone
            IMAGE:1545850 5' similar to TR:Q9Z2B4 Q9Z2B4 SMALL ESPIN. ;, mRNA
            sequence.
ACCESSION   BE138282
VERSION     BE138282.1  GI:8600782
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 467)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: ug50g06.x1
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:951198
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .467
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1545850"
                     /clone_lib="Barstead bowel MPLRB9"
                     /tissue_type="bowel"
                     /dev_stage="8 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [AATTCGGATCCTTG], digested with Not I and cloned into the
                     Not I and Eco RI sites of the modified pT7T3 vector.
                     Source irradiated bowel harvested 72 hours after
                     irradiation (1400 Gys). Library constructed by Bob
                     Barstead. "
BASE COUNT       90 a    165 c    133 g     79 t
ORIGIN


alignment_scores:
            Quality:   42.00             Length:     10
              Ratio:  4.200               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 80.000

alignment_block:
 US-08-860-232-1 x BE138282/rev   ..

 Align seg 1/1  to reverse of: BE138282  from: 1  to: 467

       2 LeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||:::|||:::|||||||||
     229 CTGCCTGAGAACACGGTTGTCAGCCCCTTG 200

seq_name: gb_est11:AI510593

seq_documentation_block:
LOCUS       AI510593      500 bp    mRNA            EST       15-MAR-2000
DEFINITION  mq40a07.y1 Barstead MPLRB1 Mus musculus cDNA clone IMAGE:581172 5'
            similar to TR:Q63618 Q63618 ESPIN. ;, mRNA sequence.
ACCESSION   AI510593
VERSION     AI510593.1  GI:4409498
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 500)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
```

```
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            This read is a RESEQUENCE of a previously sequenced mouse clone
            This read has been verified (found to hit its original self in the
            correct orientation)
            MGI:355820
            Seq primer: -40RP from Gibco
            High quality sequence stop: 490
            POLYA=No.
FEATURES              Location/Qualifiers
     source           1. .500
                      /organism="Mus musculus"
                      /strain="BALB/c"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:581172"
                      /clone_lib="Barstead MPLRB1"
                      /sex="mixed"
                      /tissue_type="Kidney"
                      /dev_stage="6 weeks"
                      /lab_host="DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      [CATGGATTCGGTACC], digested with Not I and cloned into the
                      Not I and Eco RI sites of the modified pT7T3 vector.
                      Library constructed by Bob Barstead."
BASE COUNT       96 a    174 c    145 g     85 t
ORIGIN


alignment_scores:
            Quality:   42.00             Length:     10
              Ratio:  4.200               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 80.000

alignment_block:
 US-08-860-232-1 x AI510593/rev   ..

 Align seg 1/1  to reverse of: AI510593  from: 1  to: 500

       2 LeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||:::|||:::|||||||||
     174 CTGCCTGAGAACACGGTTGTCAGCCCCTTG 145

seq_name: gb_est1:AA139619

seq_documentation_block:
LOCUS        AA139619      524 bp    mRNA            EST       13-FEB-1997
DEFINITION   mq40a07.r1 Barstead MPLRB1 Mus musculus cDNA clone IMAGE:581172 5'
             similar to TR:G402513 G402513 SMALL FORKED PROTEIN. ;, mRNA
             sequence.
ACCESSION    AA139619
VERSION      AA139619.1  GI:1701824
KEYWORDS     EST.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 524)
  AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
             Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
             Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
             Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
             Waterston,R.
  TITLE      The WashU-HHMI Mouse EST Project
  JOURNAL    Unpublished (1996)
COMMENT      Contact: Marra M/Mouse EST Project
             WashU-HHMI Mouse EST Project
             Washington University School of MedicineP
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:355820
             Possible reversed clone: similarity on wrong strand
             Seq primer: -28m13 rev2 ET from Amersham
             High quality sequence stop: 478.
FEATURES              Location/Qualifiers
     source           1. .524
                      /organism="Mus musculus"
                      /strain="BALB/c"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:581172"
                      /clone_lib="Barstead MPLRB1"
                      /sex="mixed"
                      /tissue_type="Kidney"
                      /dev_stage="6 weeks"
                      /lab_host="DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      [CATGGATTCGGTACC], digested with Not I and cloned into the
                      Not I and Eco RI sites of the modified pT7T3 vector.
                      Library constructed by Bob Barstead."
BASE COUNT      103 a    175 c    160 g     86 t
ORIGIN


alignment_scores:
            Quality:   42.00             Length:     10
              Ratio:  4.200               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 80.000

alignment_block:
 US-08-860-232-1 x AA139619/rev   ..

 Align seg 1/1  to reverse of: AA139619  from: 1  to: 524

       2 LeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||||:::|||:::|||||||||
     169 CTGCCTGAGAACACGGTTGTCAGCCCCTTG 140

seq_name: gb_est25:AW975473

seq_documentation_block:
LOCUS        AW975473      641 bp    mRNA            EST       02-JUN-2000
DEFINITION   EST387582 MAGE resequences, MAGN Homo sapiens cDNA, mRNA sequence.
ACCESSION    AW975473
VERSION      AW975473.1  GI:8166687
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 641)
  AUTHORS    Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
             ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
             Quackenbush,J.
  TITLE      Assessment of gene expression patterns in a model of colon tumor
             metastasis using a 19,200 element cDNA microarray
  JOURNAL    Unpublished (2000)
COMMENT      Contact: John Quackenbush
             The Institute for Genomic Research
             9712 Medical Center Dr., Rockville, MD 20850, USA
             Tel: 301 838 3528
             Fax: 301 838 0208
```

```
           Email: johnq@tigr.org
           Plate: 349
           Seq primer: Forward.
FEATURES             Location/Qualifiers
     source          1. .641
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGN"
                     /note="Vector: pBluescriptSKm"
BASE COUNT      219 a    117 c    124 g    181 t
ORIGIN


alignment_scores:
            Quality:   42.00             Length:     11
              Ratio:  4.200               Gaps:      0
 Percent Similarity:  90.909   Percent Identity: 72.727

alignment_block:
 US-08-860-232-1 x AW975473   ..

 Align seg 1/1  to: AW975473  from: 1  to: 641

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         |||||||||:::|||||||||:::   ||||||
     412 CTCCTCCCCCATAATAATGTCATGCTGCCTCTA 444
```

```
OM of: US-08-860-232-1 to: GenEmbl:*   out_format : pfs

Date: Dec 12, 2000  3:34 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US08860232/runat_11122000_153406_14807/app_query.fasta_1.67
 -DB=GenEmbl -QFMT=fastap -SUFFIX=lim60.rge -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=60 -USER=US08860232_@CGN1_1_3727
 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-08-860-232-1
Query length: 11
Database: GenEmbl:*
Database sequences: 1033670
Database length: -2111177393
Search time (sec): 909.340000



score_list:
Sequence              Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_pat:I22246          +   36.00   138.54   13.57     22  ! I22246 Sequence 16 from patent U
gb_pr7:HUMP5303        +   36.00   138.54   13.57     22  ! M13113 Human cellular phosphopro
gb_pat:AR060419        +   36.00   131.98   31.48     52  ! AR060419 Sequence 23 from patent
gb_pat:I22233          +   36.00   131.98   31.48     52  ! I22233 Sequence 3 from patent US
gb_pr7:HUMP53A03       +   36.00   131.98   31.48     52  ! M22883 Human phosphoprotein p53
gb_pat:A58299          +   34.00   125.79   69.62     46  ! A58299 Sequence 4 from Patent WO
gb_pat:A58300          +   34.00   125.79   69.62     46  ! A58300 Sequence 5 from Patent WO
gb_un:AX001057         -   31.00   117.41   204.00    34  ! AX001057 Sequence 8 from Patent
gb_pat:A08852          -   31.00   116.56   227.44    38  ! A08852 H.sapiens flanking sequen
gb_pat:A08853          +   31.00   116.56   227.44    38  ! A08853 H.sapiens flanking sequen
gb_pat:A98784          -   31.00   115.10   274.16    46  ! A98784 Sequence 17 from Patent W
gb_pat:AR032396        +   31.00   114.78   285.81    48  ! AR032396 Sequence 8 from patent
gb_pat:I29136          +   31.00   114.78   285.81    48  ! I29136 Sequence 8 from patent US
gb_pat:I90810          +   31.00   114.78   285.81    48  ! I90810 Sequence 8 from patent US
gb_pat:AR032467        +   31.00   114.62   291.63    49  ! AR032467 Sequence 79 from patent
gb_pat:I29207          +   31.00   114.62   291.63    49  ! I29207 Sequence 79 from patent U
gb_pat:I90881          +   31.00   114.62   291.63    49  ! I90881 Sequence 79 from patent U
gb_ro:AF005554         -   30.00   110.75   478.89    51  ! AF005554 Mus musculus T cell rec
gb_ro:AF005559         -   30.00   110.75   478.89    51  ! AF005559 Mus musculus T cell rec
gb_ro:MUSTCRBCAM       -   30.00   110.32   506.42    54  ! M80451 Mouse T-cell receptor bet
gb_pat:E13248          +   29.00   109.04   596.32    40  ! E13248 Oligonucleotide for micro
gb_pr5:HSA278639       +   28.00   104.58   1.1e+03   45  ! AJ278639 Homo sapiens partial mR
gb_ba2:CLTSTRRD        +   28.00   103.93   1.1e+03   49  ! M34687 Callothrix sp. short tand
gb_pat:I32184          +   28.00   102.52   1.4e+03   59  ! I32184 Sequence 60 from patent U
gb_pat:I34275          +   28.00   102.52   1.4e+03   59  ! I34275 Sequence 60 from patent U
gb_pat:I82480          +   28.00   102.52   1.4e+03   59  ! I82480 Sequence 60 from patent U
gb_pat:A31919          +   27.00   104.64   1.0e+03   28  ! A31919 Synthetic human alpha-IAT
gb_pat:AR011828        -   27.00   104.64   1.0e+03   28  ! AR011828 Sequence 23 from patent
gb_pat:I77151          -   27.00   104.64   1.0e+03   28  ! I77151 Sequence 23 from patent U
gb_pat:AR050016        -   27.00   104.11   1.1e+03   30  ! AR050016 Sequence 2 from patent
gb_pat:AR064964        -   27.00   104.11   1.1e+03   30  ! AR064964 Sequence 2 from patent
gb_pat:A24329          -   27.00   102.94   1.3e+03   35  ! A24329 SK70 probe. 3/1995
gb_pat:I69092          +   27.00   102.72   1.3e+03   36  ! I69092 Sequence 362 from patent
gb_pat:AR022551        -   27.00   101.92   1.5e+03   40  ! AR022551 Sequence 34 from patent
gb_pat:A26680          +   27.00   100.07   1.9e+03   51  ! A26680 Oligonucleotide primer nc
gb_pat:AR050486        +   27.00    99.22   2.1e+03   57  ! AR050486 Sequence 36 from patent
gb_pat:I44792          +   27.00    99.22   2.1e+03   57  ! I44792 Sequence 16 from patent U
gb_ro:MMTCRB235        -   26.50    98.75   2.2e+03   48  ! X70750 M.musculus mRNA for T-cel
gb_pat:A56975          -   26.00   103.64   1.2e+03   20  ! A56975 Sequence 33 from Patent W
gb_pat:E13448          -   26.00   101.94   1.5e+03   25  ! E13448 PCR primer for detecting
gb_pat:E13871          +   26.00   101.08   1.7e+03   28  ! E13871 PCR primer for gaining Tr
gb_pat:AR011864        +   26.00   100.55   1.8e+03   30  ! AR011864 Sequence 59 from pat
gb_pat:I77187          +   26.00   100.55   1.8e+03   30  ! I77187 Sequence 59 from paten
gb_pat:A08054          +   26.00   100.30   1.8e+03   31  ! A08054 Synthetic DNA fragment
gb_pat:A13208          +   26.00   100.30   1.8e+03   31  ! A13208 oligonucleotide. 12/19

seq_name: gb_pat:I22246

seq_documentation_block:
LOCUS       I22246          22 bp    DNA             PAT       07-OCT-1996
DEFINITION  Sequence 16 from patent US 5527676.
ACCESSION   I22246
VERSION     I22246.1  GI:1602600
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Vogelstein,B., Baker,S.J., Fearon,E.R. and Nigro,J.M.
  TITLE     Detection of loss of the wild-type P53 gene and kits therefor
  JOURNAL   Patent: US 5527676-A 16 18-JUN-1996;
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="unknown"
BASE COUNT        7 a      6 c      3 g      6 t
ORIGIN


alignment_scores:
            Quality:   36.00             Length:      7
              Ratio:  5.143               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x I22246   ..

 Align seg 1/1  to: I22246  from: 1  to: 22

       2 LeuProGluAsnAsnValLeu 8
         ||||||||||||||||||||
       2 CTTCCTGAAAACAACGTTCTG 22

seq_name: gb_pr7:HUMP5303

seq_documentation_block:
LOCUS       HUMP5303        22 bp    DNA             PRI       07-JAN-1995
DEFINITION  Human cellular phosphoprotein p53 gene, exon 3.
ACCESSION   M13113
VERSION     M13113.1  GI:189452
KEYWORDS    antigen; phosphoprotein; tumor antigen.
SEGMENT     3 of 11
SOURCE      Human: fetal liver DNA, clones lambda-p53-[alpha,pi];
            lymphoblastoid cell line MANN DNA (library of S.Carson), clone
            pcos53PHD; Jurkat T-cell line J6, cDNA to mRNA (library of Kataoka
            and Collins), clone p53J6K.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Lamb,P. and Crawford,L.
  TITLE     Characterization of the human p53 gene
  JOURNAL   Mol. Cell. Biol. 6 (5), 1379-1385 (1986)
  MEDLINE   87064416
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="17p13.1"
     exon            1. .22
                     /gene="TP53"
                     /note="G00-120-445"
                     /number=3
     prim_transcript <1. .>22
                     /gene="TP53"
```

```
                     /note="p53 mRNA"
BASE COUNT         7 a      6 c      3 g      6 t
ORIGIN       About 117 bp after segment 2; chromosome 17p13.


alignment_scores:
            Quality:   36.00             Length:      7
              Ratio:  5.143               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x HUMP5303   ..

 Align seg 1/1  to: HUMP5303  from: 1  to: 22

       2 LeuProGluAsnAsnValLeu 8
         ||||||||||||||||||||
       2 CTTCCTGAAAACAACGTTCTG 22

seq_name: gb_pat:AR060419

seq_documentation_block:
LOCUS        AR060419       52 bp    DNA             PAT       29-SEP-1999
DEFINITION   Sequence 23 from patent US 5840579.
ACCESSION    AR060419
VERSION      AR060419.1  GI:5986869
KEYWORDS     .
SOURCE       Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE    1  (bases 1 to 52)
  AUTHORS    Boeke,J.D. and Brachmann,R.K.
  TITLE      Nucleic acids encoding p53 mutations which suppress p53 cancer
             mutations
  JOURNAL    Patent: US 5840579-A 23 24-NOV-1998;
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="unknown"
BASE COUNT      14 a     11 c     12 g     15 t
ORIGIN


alignment_scores:
            Quality:   36.00             Length:      7
              Ratio:  5.143               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x AR060419   ..

 Align seg 1/1  to: AR060419  from: 1  to: 52

       2 LeuProGluAsnAsnValLeu 8
         ||||||||||||||||||||
      17 CTTCCTGAAAACAACGTTCTG 37

seq_name: gb_pat:I22233

seq_documentation_block:
LOCUS        I22233         52 bp    DNA             PAT       07-OCT-1996
DEFINITION   Sequence 3 from patent US 5527676.
ACCESSION    I22233
VERSION      I22233.1  GI:1602587
KEYWORDS     .
SOURCE       Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE    1  (bases 1 to 52)
  AUTHORS    Vogelstein,B., Baker,S.J., Fearon,E.R. and Nigro,J.M.
  TITLE      Detection of loss of the wild-type P53 gene and kits therefor
  JOURNAL    Patent: US 5527676-A 3 18-JUN-1996;
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="unknown"
BASE COUNT      14 a     11 c     12 g     15 t
ORIGIN


alignment_scores:
            Quality:   36.00             Length:      7
              Ratio:  5.143               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x I22233   ..

 Align seg 1/1  to: I22233  from: 1  to: 52

       2 LeuProGluAsnAsnValLeu 8
         ||||||||||||||||||||
      17 CTTCCTGAAAACAACGTTCTG 37

seq_name: gb_pr7:HUMP53A03

seq_documentation_block:
LOCUS        HUMP53A03      52 bp    DNA             PRI       08-JAN-1995
DEFINITION   Human phosphoprotein p53 gene, exon 3.
ACCESSION    M22883
VERSION      M22883.1  GI:189466
KEYWORDS     phosphoprotein p53.
SEGMENT      3 of 11
SOURCE       Homo sapiens placenta DNA.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 52)
  AUTHORS    Buchman,V.L., Chumakov,P.M., Ninkina,N.N., Samarina,O.P. and
             Georgiev,G.P.
  TITLE      A variation in the structure of the protein-coding region of the
             human p53 gene
  JOURNAL    Gene 70 (2), 245-252 (1988)
  MEDLINE    89108008
COMMENT      .
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /tissue_type="placenta"
     intron          order(M22882.1:118. .>133,<1. .15)
                     /gene="p53"
                     /number=2
     exon            16. .37
                     /gene="p53"
                     /number=3
                     /product="phosphoprotein p53"
BASE COUNT      14 a     11 c     12 g     15 t
ORIGIN       About 86 bp after segment 2.


alignment_scores:
            Quality:   36.00             Length:      7
              Ratio:  5.143               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x HUMP53A03   ..

 Align seg 1/1  to: HUMP53A03  from: 1  to: 52

       2 LeuProGluAsnAsnValLeu 8
         ||||||||||||||||||||
      17 CTTCCTGAAAACAACGTTCTG 37

seq_name: gb_pat:A58299

seq_documentation_block:
```

```
LOCUS       A58299         46 bp    DNA             PAT       05-MAR-1998
DEFINITION  Sequence 4 from Patent WO9634981.
ACCESSION   A58299
VERSION     A58299.1  GI:3713963
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Nicolaevna,M.I. and Dumas,M.E.
  TITLE     METHOD FOR THE SPECIFIC COUPLING OF THE CAP OF THE EXTREMITY 5' OF
            A FRAGMENT mRNA AND PREPARATION OF mRNA AND COMPLETE cDNA
  JOURNAL   Patent: WO 9634981-A 4 07-NOV-1996;
            GENSET (FR)
COMMENT     Other publication AU 5982996 961121
            Other publication FR 2733765 961108
            Other publication FR 2733762 961108.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       10 a      24 c       0 g      11 t       1 others
ORIGIN


alignment_scores:
            Quality:   34.00             Length:      10
              Ratio:  3.400               Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x A58299   ..

 Align seg 1/1  to: A58299  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         |||||||||:::|||::::::|||:::|||
       6 CTACTCCCATCCAATTCCACCCTAACTCCT 35

seq_name: gb_pat:A58300

seq_documentation_block:
LOCUS       A58300         46 bp    DNA             PAT       05-MAR-1998
DEFINITION  Sequence 5 from Patent WO9634981.
ACCESSION   A58300
VERSION     A58300.1  GI:3713964
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Nicolaevna,M.I. and Dumas,M.E.
  TITLE     METHOD FOR THE SPECIFIC COUPLING OF THE CAP OF THE EXTREMITY 5' OF
            A FRAGMENT mRNA AND PREPARATION OF mRNA AND COMPLETE cDNA
  JOURNAL   Patent: WO 9634981-A 5 07-NOV-1996;
            GENSET (FR)
COMMENT     Other publication AU 5982996 961121
            Other publication FR 2733765 961108
            Other publication FR 2733762 961108.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       10 a      24 c       0 g      11 t       1 others
ORIGIN


alignment_scores:
            Quality:   34.00             Length:      10
              Ratio:  3.400               Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x A58300   ..

 Align seg 1/1  to: A58300  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         |||||||||:::|||::::::|||:::|||
       6 CTACTCCCATCCAATTCCACCCTAACTCCT 35

seq_name: gb_un:AX001057

seq_documentation_block:
LOCUS       AX001057       34 bp    DNA             UNA       10-MAR-2000
DEFINITION  Sequence 8 from Patent WO9902659.
ACCESSION   AX001057
VERSION     AX001057.1  GI:7241289
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   Stempfer,G. and Schoergendorfer,K.
  TITLE     METHOD OF ALTERING THE DOMAINS OF CYCLOSPORIN SYNTHETASE TO GIVE A
            MODIFIED CYCLOSPORIN SYNTHETASE
  JOURNAL   Patent: WO 9902659-A 21-JAN-1999;
            STEMPFER GUENTHER (AT); BIOCHEMIE GMBH (AT)
FEATURES             Location/Qualifiers
     source          1. .34
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        9 a       6 c      11 g       8 t
ORIGIN


alignment_scores:
            Quality:   31.00             Length:       7
              Ratio:  4.429               Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 85.714

alignment_block:
 US-08-860-232-1 x AX001057/rev   ..

 Align seg 1/1  to reverse of: AX001057  from: 1  to: 34

       4 GluAsnAsnValLeuSerPro 10
         ||||||:::||||||||||||
      22 GAGAATTCCGTCTTAAGTCCC 2

seq_name: gb_pat:A08852

seq_documentation_block:
LOCUS       A08852         38 bp    DNA             PAT       02-SEP-1993
DEFINITION  H.sapiens flanking sequences of 33.6.
ACCESSION   A08852
VERSION     A08852.1  GI:411774
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Jeffreys,A.J.
  TITLE     Extended nucleotide sequences
  JOURNAL   Patent: EP 0370719-A 28 30-MAY-1990;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       11 a      6 c      11 g      10 t
ORIGIN


alignment_scores:
```

```
           Quality:   31.00            Length:      9
             Ratio:  3.875              Gaps:      0
 Percent Similarity:  88.889   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x A08852/rev   ..

 Align seg 1/1  to reverse of: A08852  from: 1  to: 38

       3 ProGluAsnAsnValLeuSerProLeu 11
         |||:::::|||||||   |||||||||
      33 CCAAGGTCAAATGTGAGGTCTCCTCTA 7

seq_name: gb_pat:A08853

seq_documentation_block:
LOCUS       A08853          38 bp    DNA             PAT       02-SEP-1993
DEFINITION  H.sapiens flanking sequences of 33.6, reverse complement.
ACCESSION   A08853
VERSION     A08853.1  GI:411775
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Jeffreys,A.J.
  TITLE     Extended nucleotide sequences
  JOURNAL   Patent: EP 0370719-A 29 30-MAY-1990;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       10 a     11 c      6 g     11 t
ORIGIN


alignment_scores:
           Quality:   31.00            Length:      9
             Ratio:  3.875              Gaps:      0
 Percent Similarity:  88.889   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x A08853   ..

 Align seg 1/1  to: A08853  from: 1  to: 38

       3 ProGluAsnAsnValLeuSerProLeu 11
         |||:::::|||||||   |||||||||
       6 CCAAGGTCAAATGTGAGGTCTCCTCTA 32

seq_name: gb_pat:A98784

seq_documentation_block:
LOCUS       A98784          46 bp    DNA             PAT       26-JAN-2000
DEFINITION  Sequence 17 from Patent WO9910358.
ACCESSION   A98784
VERSION     A98784.1  GI:6781805
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Hegemann,P.
  TITLE     METHOD FOR PRODUCING NUCLEIC ACID POLYMERS
  JOURNAL   Patent: WO 9910358-A 04-MAR-1999;
            HEGEMANN PETER (DE)
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        8 a      9 c     22 g      7 t
ORIGIN


alignment_scores:
           Quality:   31.00            Length:      9
             Ratio:  3.875              Gaps:      0
 Percent Similarity:  88.889   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x A98784/rev   ..

 Align seg 1/1  to reverse of: A98784  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSer 9
         |||||||||:::|||:::   ||||||
      42 CTGCTGCCCGACAACCACTACCTGTCC 16

seq_name: gb_pat:AR032396

seq_documentation_block:
LOCUS       AR032396        48 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 8 from patent US 5869241.
ACCESSION   AR032396
VERSION     AR032396.1  GI:5948001
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Method of determining DNA sequence preference of a DNA-binding
            molecule
  JOURNAL   Patent: US 5869241-A 8 09-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="unknown"
BASE COUNT       12 a     15 c      8 g     13 t
ORIGIN


alignment_scores:
           Quality:   31.00            Length:      9
             Ratio:  4.429              Gaps:      0
 Percent Similarity:  77.778   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x AR032396   ..

 Align seg 1/1  to: AR032396  from: 1  to: 48

       3 ProGluAsnAsnValLeuSerProLeu 11
         |||:::|||||   |||   ||||||
      18 CCTAATAATAACAGTCTTTGCCCTCTT 44

seq_name: gb_pat:I29136

seq_documentation_block:
LOCUS       I29136          48 bp    DNA             PAT       06-FEB-1997
DEFINITION  Sequence 8 from patent US 5578444.
ACCESSION   I29136
VERSION     I29136.1  GI:1819927
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Sequence-directed DNA-binding molecules compositions and methods
  JOURNAL   Patent: US 5578444-A 8 26-NOV-1996;
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="unknown"
BASE COUNT       12 a     15 c      8 g     13 t
```

```
ORIGIN


alignment_scores:
            Quality:   31.00             Length:      9
              Ratio:  4.429               Gaps:      0
 Percent Similarity:  77.778   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x I29136   ..

 Align seg 1/1  to: I29136  from: 1  to: 48

       3 ProGluAsnAsnValLeuSerProLeu 11
         |||:::||||||   |||   ||||||
      18 CCTAATAATAACAGTCTTTGCCCTCTT 44

seq_name: gb_pat:I90810

seq_documentation_block:
LOCUS       I90810         48 bp    DNA             PAT       01-DEC-1998
DEFINITION  Sequence 8 from patent US 5726014.
ACCESSION   I90810
VERSION     I90810.1  GI:3935280
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.
  TITLE     Screening assay for the detection of DNA-binding molecules
  JOURNAL   Patent: US 5726014-A 8 10-MAR-1998;
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="unknown"
BASE COUNT       12 a     15 c      8 g     13 t
ORIGIN


alignment_scores:
            Quality:   31.00             Length:      9
              Ratio:  4.429               Gaps:      0
 Percent Similarity:  77.778   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x I90810   ..

 Align seg 1/1  to: I90810  from: 1  to: 48

       3 ProGluAsnAsnValLeuSerProLeu 11
         |||:::||||||   |||   ||||||
      18 CCTAATAATAACAGTCTTTGCCCTCTT 44

seq_name: gb_pat:AR032467

seq_documentation_block:
LOCUS       AR032467       49 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 79 from patent US 5869241.
ACCESSION   AR032467
VERSION     AR032467.1  GI:5948072
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Method of determining DNA sequence preference of a DNA-binding
            molecule
  JOURNAL   Patent: US 5869241-A 79 09-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="unknown"
BASE COUNT       13 a     15 c      8 g     13 t
```

```
ORIGIN


alignment_scores:
            Quality:   31.00             Length:      9
              Ratio:  4.429               Gaps:      0
 Percent Similarity:  77.778   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x AR032467   ..

 Align seg 1/1  to: AR032467  from: 1  to: 49

       3 ProGluAsnAsnValLeuSerProLeu 11
         |||:::||||||   |||   ||||||
      19 CCTAATAATAACAGTCTTTGCCCTCTT 45
```

```
OM of: US-08-860-232-1 to: N_Geneseq_36:*   out_format : pfs

Date: Dec 12, 2000  3:37 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US08860232/runat_11122000_153407_14843/app_query.fasta_1.67
 -DB=N_Geneseq_36 -QFMT=fastap -SUFFIX=lim60.rng -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=60 -USER=US08860232_@CGN1_1_108
 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-08-860-232-1
Query length: 11
Database: N_Geneseq_36:*
Database sequences: 480022
Database length: 187831343
Search time (sec): 71.090000



score_list:
Sequence            Strd Orig    ZScore    EScore Len  ! Documentation  ..
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:Z40750 +    34.00  127.94  24.45     46  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X88171 +    34.00  127.94  24.45     46  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X97544 +    34.00  127.94  24.45     46  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X19963 +    34.00  127.94  24.45     46  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X51429 +    34.00  127.94  24.45     46  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X39410 +    34.00  127.94  24.45     46  ! Hum
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X40408 +    34.00  127.94  24.45     46  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X51757 +    34.00  127.94  24.45     46  ! Unc
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X26652 +    34.00  127.94  24.45     46  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X41349 +    34.00  127.94  24.45     46  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X30063 +    34.00  127.94  24.45     46  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T43581 +    34.00  127.75  25.08     47  ! 5'-
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:Z40749 +    34.00  127.75  25.08     47  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X88170 +    34.00  127.75  25.08     47  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X97543 +    34.00  127.75  25.08     47  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X19962 +    34.00  127.75  25.08     47  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X51428 +    34.00  127.75  25.08     47  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X39409 +    34.00  127.75  25.08     47  ! Hum
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X40407 +    34.00  127.75  25.08     47  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X51756 +    34.00  127.75  25.08     47  ! Cap
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X26651 +    34.00  127.75  25.08     47  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X41348 +    34.00  127.75  25.08     47  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X30062 +    34.00  127.75  25.08     47  ! Oli
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X15797 -    31.00  120.34  64.88     34  ! PCR
/SIDS6/gcgdata/geneseq/geneseqn/NA1995.DAT:Q84796 -    31.00  119.81  69.39     36  ! Spi
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X22925 -    31.00  117.56  92.59     46  ! DEI
/SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:Q69258 +    31.00  117.17  97.35     48  ! Hum
/SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:T63720 +    31.00  117.17  97.35     48  ! Hum
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X17008 +    31.00  117.17  97.35     48  ! Tes
/SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:Q69329 +    31.00  116.98  99.74     49  ! Hum
/SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:T63791 +    31.00  116.98  99.74     49  ! Hum
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X17079 +    31.00  116.98  99.74     49  ! Tes
/SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:Q62352 +    30.00  118.02  87.27     30  ! PCR
/SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:T78971 -    29.00  118.28  84.41     20  ! Pri
/SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:T79520 +    29.00  113.69  152.16    33  ! Fil
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X26307 +    29.00  113.69  152.16    33  ! PCR
/SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:Z23458 +    29.00  113.42  157.60    34  ! Pla
/SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:Z55429 +    29.00  112.89  168.57    36  ! Nei
/SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:T78191 +    29.00  111.92  190.82    40  ! Mic
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V36363 -    29.00  110.64  224.93    46  ! Tem
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V23543 +    28.00  111.10  212.00    30  ! Mou
/SIDS6/gcgdata/geneseq/geneseqn/NA1992.DAT:Q28801 +    28.00  109.69  254.17    35  !
/SIDS6/gcgdata/geneseq/geneseqn/NA1989.DAT:N91674 -    28.00  106.24  395.84    51  !
/SIDS6/gcgdata/geneseq/geneseqn/NA1995.DAT:Q80186 +    28.00  104.90  469.88    59  !
/SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:V04982 +    28.00  104.90  469.88    59  !

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:Z40750

seq_documentation_block:
ID   Z40750 standard; RNA; 46 BP.
XX
AC   Z40750;
XX
DT   18-JAN-2000  (first entry)
XX
DE   Oligonucleotide -Cap for Secreted protein EST isolation.
XX
KW   PCR primer; secreted protein; fingerprint identification technique;
KW   chromosome mapping; human; hereditary disease; diagnosis; cancer;
KW   hyperlipidaemia; cardiovascular; neurodegenerative disorder; therapy;
KW   autoimmune disease; rheumatic disease; embryogenic disorder; myopathy;
KW   renal injury; amino aciduria; hypoglycaemia; male rat infertility;
KW   hypertension; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9940189-A2.
XX
PD   12-AUG-1999.
XX
PF   09-FEB-1999;   99WO-IB00282.
XX
PR   09-FEB-1998;   98US-0074121.
PR   13-APR-1998;   98US-0081563.
PR   10-AUG-1998;   98US-0096116.
PR   04-SEP-1998;   98US-0099273.
XX
PA   (GEST ) GENSET.
XX
PI   Bougueleret L,  Duclert A,  Dumas Milne Edwards J;
XX
DR   WPI; 1999-600966/51.
XX
PT   Extended cDNAs useful for expressing secreted proteins and to obtain
PT   specific antibodies -
XX
PS   Example 2; Page 12; 244pp; English.
XX
CC   This sequence represents a PCR primer used within the course of the
CC   invention. The invention relates to 70 nucleic acids encoding human
CC   secreted proteins. The extended cDNAs (or genomic DNAs obtainable from
CC   them) may be used to prepare PCR primers and probes. These are useful for
CC   forensic matching or positive identification by DNA sequencing. They may
CC   also be used in alternative fingerprint identification techniques.
CC   Antibodies against the proteins encoded by the extended cDNAs are useful
CC   in identification of tissue types or cell species, as well as identifying
CC   tissue specific soluble proteins. The sequences can be used for
CC   chromosome mapping and identification of genes associated with hereditary
CC   diseases or drug response. signal sequences from the cDNAs can be used in
CC   construction of secretion vectors. Other sequences derived from the
CC   extended cDNAs can be used to clone upstream genomic DNA sequences
CC   including promoters. This is in turn useful for identifying proteins that
CC   interact with promoter sequences. Some of the proteins may be useful in
CC   diagnosing and treating several disorders including, but not limited to:
CC   cancer, hyperlipidaemia, cardiovascular and neurodegenerative disorders,
CC   autoimmune diseases, and rheumatic diseases, embryogenic disorders,
CC   hypertension, renal injury, amino acidurias, hypoglycaemia, male rat
CC   infertility and myopathies.
XX
SQ   Sequence 46 BP; 10 A; 24 C; 1 G; 11 U; 0 other;


alignment_scores:
```

```
            Quality:   34.00             Length:     10
              Ratio:  3.400               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x Z40750   ..

 Align seg 1/1  to: Z40750  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         ||||||||||:::|||:::::::|||:::|||
       6 CUACUCCCAUCCAAUUCCACCCUAACUCCU 35

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X88171

seq_documentation_block:
ID   X88171 standard; RNA; 46 BP.
XX
AC   X88171;
XX
DT   23-SEP-1999  (first entry)
XX
DE   Oligoribonucleotide Cap5'-ppp.
XX
KW   Secreted protein; human; cytostatic; thrombotic; osteopathic; forensic;
KW   diagnostic; gene therapy; chromosome mapping; secretion vector; primer;
KW   ss.
XX
OS   Synthetic.
XX
PN   WO9925825-A2.
XX
PD   27-MAY-1999.
XX
PF   13-NOV-1998;   98WO-IB01862.
XX
PR   04-SEP-1998;   98US-0099273.
PR   13-NOV-1997;   97US-0066677.
PR   17-DEC-1997;   97US-0069957.
PR   09-FEB-1998;   98US-0074121.
PR   13-APR-1998;   98US-0081563.
PR   10-AUG-1998;   98US-0096116.
XX
PA   (GEST ) GENSET.
XX
PI   Bougueleret L,  Duclert A,  Dumas Milne Edwards J;
XX
DR   WPI; 1999-347472/29.
XX
PT   Extended cDNAs encoding secreted proteins
XX
PS   Example 2; Page 131; 307pp; English.
XX
CC   This invention describes novel nucleic acid sequences of extended cDNAs
CC   (see X97813-X97906) which encode human secreted proteins (see
CC   Y36129-Y36222) and which have cytostatic, thrombotic and osteopathic
CC   activity. The extended cDNAs can be used to express secreted proteins
CC   or parts of them or to obtain antibodies capable of binding to the
CC   secreted proteins. They may also be used in diagnostic, forensic,
CC   gene therapy and chromosome mapping procedures. Uses also include design
CC   of expression vectors and secretion vectors. This sequence represents
CC   an oligoribonucleotide primer used in the method of the invention.
XX
SQ   Sequence 46 BP; 10 A; 24 C; 1 G; 11 U; 0 other;


alignment_scores:
            Quality:   34.00             Length:     10
              Ratio:  3.400               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x X88171   ..

 Align seg 1/1  to: X88171  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         ||||||||||:::|||:::::::|||:::|||
       6 CUACUCCCAUCCAAUUCCACCCUAACUCCU 35

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X97544

seq_documentation_block:
ID   X97544 standard; RNA; 46 BP.
XX
AC   X97544;
XX
DT   13-SEP-1999  (first entry)
XX
DE   Oligonucleotide Cap- for secreted protein coding sequence isolation.
XX
KW   Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW   cellular differentiation; immune system regulator; anti-inflammatory;
KW   haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW   reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW   genetic disease; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9931236-A2.
XX
PD   24-JUN-1999.
XX
PF   17-DEC-1998;   98WO-IB02122.
XX
PR   10-AUG-1998;   98US-0096116.
PR   17-DEC-1997;   97US-0069957.
PR   09-FEB-1998;   98US-0074121.
PR   13-APR-1998;   98US-0081563.
XX
PA   (GEST ) GENSET.
XX
PI   Bougueleret L,  Duclert A,  Dumas Milne Edwards J;
XX
DR   WPI; 1999-385906/32.
XX
PT   New isolated human secreted proteins
XX
PS   Example 2; Page 12; 516pp; English.
XX
CC   This sequence represents an oligonucleotide used to isolate the extended
CC   human secreted protein coding sequences of the invention. The secreted
CC   proteins can be used in treating or controlling a variety of human
CC   conditions. The secreted proteins may act as cytokines or may affect
CC   cellular proliferation or differentiation or may act as immune system
CC   regulators, haematopoiesis regulators, tissue growth regulators,
CC   regulators of reproductive hormones or cell movement or have
CC   chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or tumour
CC   inhibition activity. The DNAs can be used in forensic procedures to
CC   identify individuals or in diagnostic procedures to identify individuals
CC   having genetic diseases resulting from abnormal expression of the genes
CC   corresponding to the extended cDNAs. They are also useful for
CC   constructing a high resolution map of the human chromosomes. They can
CC   also be used for gene therapy to control or treat genetic diseases.
XX
SQ   Sequence 46 BP; 10 A; 24 C; 1 G; 11 U; 0 other;


alignment_scores:
            Quality:   34.00             Length:     10
              Ratio:  3.400               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x X97544   ..
```

```
 Align seg 1/1  to: X97544  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         |||||||||:::|||::::::|||:::|||
       6 CUACUCCCAUCCAAUUCCACCCUAACUCCU 35

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X19963

seq_documentation_block:
ID   X19963 standard; RNA; 46 BP.
XX
AC   X19963;
XX
DT   16-JUN-1999  (first entry)
XX
DE   Oligoribonucleotide -Cap SEQ ID NO:2.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition; ss.
XX
OS   Synthetic.
XX
PN   WO9906439-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB01233.
XX
PR   01-AUG-1997;   97US-0904468.
XX
PA   (GEST ) GENSET.
XX
PI   Duclert A,  Dumas Milne Edwards J,  Lacroix B;
XX
DR   WPI; 1999-153700/13.
XX
PT   New nucleic acids encoding human secreted proteins - obtained from
PT   cDNA libraries derived from liver, lung, large intestine, colon,
PT   thyroid and pancreas tissue
XX
PS   Example 2; Page 15; 398pp; English.
XX
CC   X40251 to X40397 represent 5' expressed sequence tags (ESTs) for human
CC   secreted proteins, and encode the proteins given in Y11533 to Y11679,
CC   respectively. The proteins given represent the signal peptide and an
CC   N-terminal fragment of a secreted protein. The nucleic acid sequences
CC   can be used for producing secreted human gene products. They can also
CC   be used to develop products for diagnosis and therapy. The proteins
CC   obtained may have cytokine activity, cell proliferation/differentiation
CC   activity, haematopoiesis regulating activity, tissue growth regulating
CC   activity, reproductive hormone regulating activity, chemotactic/
CC   chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC   ligand activity, anti-inflammatory activity, tumour inhibition activity
CC   or other activities. The products can be used in forensic, gene therapy
CC   and chromosome mapping procedures. The sequences can also be used for
CC   obtaining corresponding promoter sequences. The nucleic acids encoding
CC   the signal peptide can be used for directing extracellular secretion of
CC   a polypeptide or the insertion of a polypeptide into a membrane, or
CC   importing a polypeptide into a cell. The present sequence represents an
CC   oligoribonucleotide used in an example from the present invention.
XX
SQ   Sequence 46 BP; 10 A; 24 C; 1 G; 11 U; 0 other;


alignment_scores:
            Quality:   34.00            Length:     10
              Ratio:  3.400              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000
```

```
alignment_block:
 US-08-860-232-1 x X19963   ..

 Align seg 1/1  to: X19963  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         |||||||||:::|||::::::|||:::|||
       6 CUACUCCCAUCCAAUUCCACCCUAACUCCU 35

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X51429

seq_documentation_block:
ID   X51429 standard; RNA; 46 BP.
XX
AC   X51429;
XX
DT   21-JUN-1999  (first entry)
XX
DE   Oligonucleotide -Cap.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition; ss.
XX
OS   Synthetic.
XX
PN   WO9906549-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB01231.
XX
PR   01-AUG-1997;   97US-0905279.
XX
PA   (GEST ) GENSET.
XX
PI   Duclert A,  Dumas Milne Edwards J,  Lacroix B;
XX
DR   WPI; 1999-153779/13.
XX
PT   New nucleic acids encoding human secreted proteins - obtained from
PT   cDNA libraries derived from testis, ovary, uterus and spleen tissue
XX
PS   Example 2; Page 144; 522pp; English.
XX
CC   X51459 to X51691 represent 5' expressed sequence tags (ESTs) for human
CC   secreted proteins, and encode the proteins given in Y12681 to Y12913,
CC   respectively. The proteins given represent the signal peptide and an
CC   N-terminal fragment of a secreted protein. The nucleic acid sequences
CC   can be used for producing secreted human gene products. They can also
CC   be used to develop products for diagnosis and therapy. The proteins
CC   obtained may have cytokine activity, cell proliferation/differentiation
CC   activity, haematopoiesis regulating activity, tissue growth regulating
CC   activity, reproductive hormone regulating activity, chemotactic/
CC   chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC   ligand activity, anti-inflammatory activity, tumour inhibition activity
CC   or other activities. The products can be used in forensic, gene therapy
CC   and chromosome mapping procedures. The sequences can also be used for
CC   obtaining corresponding promoter sequences. The nucleic acids encoding
CC   the signal peptide can be used for directing extracellular secretion of
CC   a polypeptide or the insertion of a polypeptide into a membrane, or
CC   importing a polypeptide into a cell. This sequence represents an
CC   oligonucleotide -Cap, used in the isolation of the 5' EST sequences of
CC   the invention.
XX
SQ   Sequence 46 BP; 10 A; 24 C; 1 G; 11 U; 0 other;


alignment_scores:
```

```
           Quality:   34.00              Length:     10
             Ratio:  3.400                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x X51429   ..

 Align seg 1/1  to: X51429  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         |||||||||:::|||::::::|||:::|||
       6 CUACUCCCAUCCAAUUCCACCCUAACUCCU 35

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X39410

seq_documentation_block:
ID   X39410 standard; RNA; 46 BP.
XX
AC   X39410;
XX
DT   21-JUN-1999  (first entry)
XX
DE   Human secreted protein 5' EST detecting oligoribonucleotide #2.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition; ss.
XX
OS   Synthetic.
XX
PN   WO9906551-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB01235.
XX
PR   01-AUG-1997;   97US-0905133.
XX
PA   (GEST ) GENSET.
XX
PI   Duclert A,  Dumas Milne Edwards J,  Lacroix B;
XX
DR   WPI; 1999-153781/13.
XX
PT   New nucleic acids encoding human secreted - proteins obtained from
PT   cDNA libraries prepared from substantia nigra, cerebellum, surrenals
PT   and fetal brain tissue
XX
PS   Example 2; Page 15; 434pp; English.
XX
CC   This invention describes 5' expressed sequence tags (ESTs) represented in
CC   X39440 to X39597, which encode the human secreted proteins represented in
CC   Y11374-Y11531. The proteins given represent the signal peptide and an
CC   N-terminal fragment of a secreted protein. The nucleic acid sequences
CC   can be used for producing secreted human gene products. They can also
CC   be used to develop products for diagnosis and therapy. The proteins
CC   obtained may have cytokine activity, cell proliferation/differentiation
CC   activity, haematopoiesis regulating activity, tissue growth regulating
CC   activity, reproductive hormone regulating activity, chemotactic/
CC   chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC   ligand activity, anti-inflammatory activity, tumour inhibition activity
CC   or other activities. The products can be used in forensic, gene therapy
CC   and chromosome mapping procedures. The sequences can also be used for
CC   obtaining corresponding promoter sequences. The nucleic acids encoding
CC   the signal peptide can be used for directing extracellular secretion of
CC   a polypeptide or the insertion of a polypeptide into a membrane, or
CC   importing a polypeptide into a cell. This sequence represents an
CC   oligoribonucleotide which is used in the method of the invention.
XX
SQ   Sequence 46 BP; 10 A; 24 C; 1 G; 11 U; 0 other;

alignment_scores:
           Quality:   34.00              Length:     10
             Ratio:  3.400                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x X39410   ..

 Align seg 1/1  to: X39410  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         |||||||||:::|||::::::|||:::|||
       6 CUACUCCCAUCCAAUUCCACCCUAACUCCU 35

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X40408

seq_documentation_block:
ID   X40408 standard; RNA; 46 BP.
XX
AC   X40408;
XX
DT   18-JUN-1999  (first entry)
XX
DE   Oligoribonucleotide -Cap.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide; prostate;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9906550-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB01232.
XX
PR   01-AUG-1997;   97US-0905144.
XX
PA   (GEST ) GENSET.
XX
PI   Duclert A,  Dumas Milne Edwards J,  Lacroix B;
XX
DR   WPI; 1999-153780/13.
XX
PT   New isolated prostate-derived nucleic acids - used to develop
PT   products which may have cytokine, immune regulatory, haematopoiesis
PT   regulating, anti-inflammatory or tumour inhibition activity
XX
PS   Example 2; Page 15; 675pp; English.
XX
CC   X40438 to X40715 represent 5' expressed sequence tags (ESTs) for human
CC   secreted proteins expressed in prostate, and encode the proteins given in
CC   Y11716 to Y11993 respectively. The proteins given represent the signal
CC   peptide and an N-terminal fragment of a secreted protein. The nucleic
CC   acid sequences can be used for producing secreted human gene products.
CC   They can also be used to develop products for diagnosis and therapy. The
CC   proteins obtained may have cytokine activity, cell proliferation and
CC   differentiation activity, haematopoiesis regulating activity, tissue
CC   growth regulating activity, reproductive hormone regulating activity,
CC   chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC   receptor/ligand activity, anti-inflammatory activity, tumour inhibition
CC   activity or other activities. The products can be used in forensic, gene
CC   therapy and chromosome mapping procedures. The sequences can also be used
CC   for obtaining corresponding promoter sequences. The nucleic acids
CC   encoding the signal peptides can be used for directing extracellular
CC   secretion of a polypeptide or the insertion of a polypeptide into a
```

```
CC   membrane, or importing a polypeptide into a cell.
XX
SQ   Sequence 46 BP; 10 A; 24 C; 1 G; 11 U; 0 other;


alignment_scores:
            Quality:   34.00              Length:     10
              Ratio:  3.400                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x X40408   ..

 Align seg 1/1  to: X40408  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         ||||||||||:::|||::::::|||:::|||
       6 CUACUCCCAUCCAAUUCCACCCUAACUCCU 35

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X51757

seq_documentation_block:
ID   X51757 standard; RNA; 46 BP.
XX
AC   X51757;
XX
DT   22-JUN-1999  (first entry)
XX
DE   Uncapped mRNA for EST sequence of human secreted protein.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition; ss.
XX
OS   Homo sapiens.
XX
PN   WO9906552-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB01236.
XX
PR   01-AUG-1997;   97US-0905223.
XX
PA   (GEST ) GENSET.
XX
PI   Duclert A,  Dumas Milne Edwards J,  Lacroix B;
XX
DR   WPI; 1999-153782/13.
XX
PT   New isolated brain-derived nucleic acids - used to develop products
PT   which may have cytokine, immune, regulatory, haematopoiesis
PT   regulating, anti-inflammatory or tumour inhibition activity
XX
PS   Example 2; Page 15; 577pp; English.
XX
CC   X51787 to X52019 represent 5' expressed sequence tags (ESTs) for human
CC   secreted proteins, and encode the proteins given in Y12987 to Y13219,
CC   respectively. The proteins given represent the signal peptide and an
CC   N-terminal fragment of a secreted protein. The nucleic acid sequences
CC   can be used for producing secreted human gene products. They can also
CC   be used to develop products for diagnosis and therapy. The proteins
CC   obtained may have cytokine activity, cell proliferation/differentiation
CC   activity, haematopoiesis regulating activity, tissue growth regulating
CC   activity, reproductive hormone regulating activity, chemotactic/
CC   chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC   ligand activity, anti-inflammatory activity, tumour inhibition activity
CC   or other activities. The products can be used in forensic, gene therapy
CC   and chromosome mapping procedures. The sequences can also be used for
CC   obtaining corresponding promoter sequences. The nucleic acids encoding
CC   the signal peptide can be used for directing extracellular secretion of
CC   a polypeptide or the insertion of a polypeptide into a membrane, or
CC   importing a polypeptide into a cell.
CC   This sequence was used in a method to isolate the 5' ESTs of the genes
CC   encoding the human secreted proteins of the invention.
XX
SQ   Sequence 46 BP; 10 A; 24 C; 1 G; 11 U; 0 other;


alignment_scores:
            Quality:   34.00              Length:     10
              Ratio:  3.400                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x X51757   ..

 Align seg 1/1  to: X51757  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         ||||||||||:::|||::::::|||:::|||
       6 CUACUCCCAUCCAAUUCCACCCUAACUCCU 35

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X26652

seq_documentation_block:
ID   X26652 standard; RNA; 46 BP.
XX
AC   X26652;
XX
DT   18-JUN-1999  (first entry)
XX
DE   Oligoribonucleotide used to identify 5' EST sequences.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition; ss.
XX
OS   Synthetic;
XX
PN   WO9906554-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB01238.
XX
PR   01-AUG-1997;   97US-0905134.
XX
PA   (GEST ) GENSET.
XX
PI   Duclert A,  Dumas Milne Edwards J,  Lacroix B;
XX
DR   WPI; 1999-153784/13.
XX
PT   New nucleic acids encoding human secreted proteins - obtained from
PT   cDNA libraries prepared from kidney, fetal kidney, dystrophic
PT   muscle, muscle and heart tissue
XX
PS   Example 2; Page 15; 622pp; English.
XX
CC   The specification describes 5' expressed sequence tags (ESTs, see
CC   X40826-X41093) for human secreted proteins (see Y01602 and Y11994-
CC   Y12260). The proteins given represent the signal peptide and an
CC   N-terminal fragment of a secreted protein. The nucleic acid sequences
CC   can be used for producing secreted human gene products. They can also
CC   be used to develop products for diagnosis and therapy. The proteins
CC   obtained may have cytokine activity, cell proliferation/differentiation
CC   activity, haematopoiesis regulating activity, tissue growth regulating
CC   activity, reproductive hormone regulating activity, chemotactic/
CC   chemokinetic activity, haemostatic and thrombolytic activity, receptor/
```

```
CC   ligand activity, anti-inflammatory activity, tumour inhibition activity
CC   or other activities. The products can be used in forensic, gene therapy
CC   and chromosome mapping procedures. The sequences can also be used for
CC   obtaining corresponding promoter sequences. The nucleic acids encoding
CC   the signal peptide can be used for directing extracellular secretion of
CC   a polypeptide or the insertion of a polypeptide into a membrane, or
CC   importing a polypeptide into a cell. The present sequence is used
CC   in the course of the invention.
XX
SQ   Sequence 46 BP; 10 A; 24 C; 1 G; 11 U; 0 other;


alignment_scores:
            Quality:   34.00              Length:     10
              Ratio:  3.400                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x X26652   ..

 Align seg 1/1  to: X26652  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         ||||||||||:::|||:::::::|||:::|||
       6 CUACUCCCAUCCAAUUCCACCCUAACUCCU 35

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X41349

seq_documentation_block:
ID   X41349 standard; RNA; 46 BP.
XX
AC   X41349;
XX
DT   22-JUN-1999  (first entry)
XX
DE   Oligoribonucleotide, SEQ ID NO: 2 from WO9906553.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; antiinflammatory; tumour inhibition; antitumour; ss.
XX
OS   Homo sapiens.
XX
PN   WO9906553-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB01237.
XX
PR   01-AUG-1997;   97US-0905051.
XX
PA   (GEST ) GENSET.
XX
PI   Duclert A,  Dumas Milne Edwards J,  Lacroix B;
XX
DR   WPI; 1999-153783/13.
XX
PT   New nucleic acids encoding human secreted proteins - obtained from
PT   cDNA libraries derived from umbilical cord, lymph ganglia,
PT   lymphocytes and placental tissue
XX
PS   Example 2; Page 15; 411pp; English.
XX
CC   The patent relates to sequences of 5' ESTs derived from mRNAs
CC   encoding secreted proteins. The nucleic acid sequences can
CC   be used for producing secreted human gene products. They can also
CC   be used to develop products for diagnosis and therapy. The proteins
CC   obtained may have cytokine activity, cell proliferation/differentiation
CC   activity, haematopoiesis regulating activity, tissue growth regulating
CC   activity, reproductive hormone regulating activity, chemotactic/
CC   chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC   ligand activity, antiinflammatory activity, tumour inhibition activity
CC   or other activities. The products can be used in forensic, gene therapy
CC   and chromosome mapping procedures. The sequences can also be used for
CC   obtaining corresponding promoter sequences. The nucleic acids encoding
CC   the signal peptide can be used for directing extracellular secretion of
CC   a polypeptide or the insertion of a polypeptide into a membrane, or
CC   importing a polypeptide into a cell.
XX
SQ   Sequence 46 BP; 10 A; 24 C; 1 G; 11 U; 0 other;


alignment_scores:
            Quality:   34.00              Length:     10
              Ratio:  3.400                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x X41349   ..

 Align seg 1/1  to: X41349  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         ||||||||||:::|||:::::::|||:::|||
       6 CUACUCCCAUCCAAUUCCACCCUAACUCCU 35

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X30063

seq_documentation_block:
ID   X30063 standard; RNA; 46 BP.
XX
AC   X30063;
XX
DT   17-JUN-1999  (first entry)
XX
DE   Oligoribonucleotide -Cap SEQ ID NO:2.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition; ss.
XX
OS   Synthetic.
XX
PN   WO9906548-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB01222.
XX
PR   01-AUG-1997;   97US-0905135.
XX
PA   (GEST ) GENSET.
XX
PI   Duclert A,  Dumas Milne Edwards J,  Lacroix B;
XX
DR   WPI; 1999-153778/13.
XX
PT   New nucleic acids encoding human secreted proteins - obtained from
PT   cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
PT   kidney, lung, umbilical cord, placenta and colon tissue
XX
PS   Example 2; Page 15; 824pp; English.
XX
CC   X41094 to X41347 represent 5' expressed sequence tags (ESTs) for human
CC   secreted proteins, and encode the proteins given in Y12261 to Y12514,
CC   respectively. The proteins given represent the signal peptide and an
CC   N-terminal fragment of a secreted protein. The nucleic acid sequences
CC   can be used for producing secreted human gene products. They can also
CC   be used to develop products for diagnosis and therapy. The proteins
CC   obtained may have cytokine activity, cell proliferation/differentiation
```

```
CC   activity, haematopoiesis regulating activity, tissue growth regulating
CC   activity, reproductive hormone regulating activity, chemotactic/
CC   chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC   ligand activity, anti-inflammatory activity, tumour inhibition activity
CC   or other activities. The products can be used in forensic, gene therapy
CC   and chromosome mapping procedures. The sequences can also be used for
CC   obtaining corresponding promoter sequences. The nucleic acids encoding
CC   the signal peptide can be used for directing extracellular secretion of
CC   a polypeptide or the insertion of a polypeptide into a membrane, or
CC   importing a polypeptide into a cell. The present sequence represents an
CC   oligoribonucleotide used in an example from the present invention.
XX
SQ   Sequence 46 BP; 10 A; 24 C; 1 G; 11 U; 0 other;


alignment_scores:
            Quality:   34.00              Length:     10
              Ratio:  3.400                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x X30063   ..

 Align seg 1/1  to: X30063  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         ||||||||||:::|||::::::|||:::|||
       6 CUACUCCCAUCCAAUUCCACCCUAACUCCU 35

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:T43581

seq_documentation_block:
ID   T43581 standard; mRNA; 47 BP.
XX
AC   T43581;
XX
DT   04-AUG-1997  (first entry)
XX
DE   5'-capped mRNA for coupling to biotin label via amino linker.
XX
KW   Messenger RNA; guanosine 5'-cap; label; immobilisation; capture;
KW   polymerase chain reaction; transcription template; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1
FT                   /*tag= a
FT                   /mod_base= m7g
FT                   /function= cap
FT                   /note= "linked to adjacent nucleotide via
FT                          (5')ppp(5') linkage"
XX
PN   WO9634981-A2.
XX
PD   07-NOV-1996.
XX
PF   29-APR-1996;   96WO-FR00651.
XX
PR   03-AUG-1995;   95FR-0009467.
PR   02-MAY-1995;   95FR-0005221.
XX
PA   (GENQ ) GENSET.
XX
PI   Dumas Milne Edwards JG,  Nicolaevna Merenkova I;
XX
DR   WPI; 1996-506181/50.
XX
PT   Specific coupling of the 5' cap of mRNA to amino-functionalised cpd.
PT   - by eliminating 3' diol, oxidn. of cap diol to aldehyde and
PT   reaction with amine, e.g. for isolation of complete RNA, labelling
PT   etc.
XX
PS   Example 1; Page 20; 49pp; French.
XX
CC   The first step in a new method for specifically coupling the cap of
CC   the 5'-end of a eukaryotic mRNA to an amino-functionalised compound
CC   involves specifically modifying the 3'-end of the mRNA so that the
CC   last base no longer contains OH groups at the 2' and 3' positions.
CC   Then, the 2', 3'-cis diol of the methyl guanosine residue at the 5'-end
CC   can be oxidised to form a 2',3'-dialdehyde which is ultimately coupled
CC   with the amino group of the amino-functionalised compound. The method
CC   is used to label specifically at the 5'-cap, to isolate the 5'-end
CC   of mRNA in a sample, to produce the 3'-end of cDNA, to produce double
CC   stranded cDNA complementary to the 5'-end of mRNA or to capture mRNA-
CC   binding proteins. In a specific example of the coupling method, a
CC   double-stranded template was prepared by PCR amplification using a
CC   5'-primer containing the T7 RNA polymerase promoter (T43579) and
CC   a 3'-primer (T43580). When the template was transcribed in the
CC   presence of cap analogue m7G(5')ppp(5')G, a capped RNA transcript
CC   having the present sequence was produced. The 5'-capped mRNA was
CC   coupled to a biotin label via a hydrazine linker.
XX
SQ   Sequence 47 BP; 10 A; 24 C; 2 G; 11 U; 0 other;


alignment_scores:
            Quality:   34.00              Length:     10
              Ratio:  3.400                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x T43581   ..

 Align seg 1/1  to: T43581  from: 1  to: 47

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         ||||||||||:::|||::::::|||:::|||
       7 CUACUCCCAUCCAAUUCCACCCUAACUCCU 36

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:Z40749

seq_documentation_block:
ID   Z40749 standard; RNA; 47 BP.
XX
AC   Z40749;
XX
DT   18-JAN-2000  (first entry)
XX
DE   Oligonucleotide +Cap for Secreted protein EST isolation.
XX
KW   PCR primer; secreted protein; fingerprint identification technique;
KW   chromosome mapping; human; hereditary disease; diagnosis; cancer;
KW   hyperlipidaemia; cardiovascular; neurodegenerative disorder; therapy;
KW   autoimmune disease; rheumatic disease; embryogenic disorder; myopathy;
KW   renal injury; amino aciduria; hypoglycaemia; male rat infertility;
KW   hypertension; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9940189-A2.
XX
PD   12-AUG-1999.
XX
PF   09-FEB-1999;   99WO-IB00282.
XX
PR   09-FEB-1998;   98US-0074121.
PR   13-APR-1998;   98US-0081563.
PR   10-AUG-1998;   98US-0096116.
PR   04-SEP-1998;   98US-0099273.
XX
PA   (GEST ) GENSET.
XX
PI   Bougueleret L,  Duclert A,  Dumas Milne Edwards J;
XX
```

```
DR   WPI; 1999-600966/51.
XX
PT   Extended cDNAs useful for expressing secreted proteins and to obtain
PT   specific antibodies -
XX
PS   Example 2; Page 12; 244pp; English.
XX
CC   This sequence represents a PCR primer used within the course of the
CC   invention. The invention relates to 70 nucleic acids encoding human
CC   secreted proteins. The extended cDNAs (or genomic DNAs obtainable from
CC   them) may be used to prepare PCR primers and probes. These are useful for
CC   forensic matching or positive identification by DNA sequencing. They may
CC   also be used in alternative fingerprint identification techniques.
CC   Antibodies against the proteins encoded by the extended cDNAs are useful
CC   in identification of tissue types or cell species, as well as identifying
CC   tissue specific soluble proteins. The sequences can be used for
CC   chromosome mapping and identification of genes associated with hereditary
CC   diseases or drug response. signal sequences from the cDNAs can be used in
CC   construction of secretion vectors. Other sequences derived from the
CC   extended cDNAs can be used to clone upstream genomic DNA sequences
CC   including promoters. This is in turn useful for identifying proteins that
CC   interact with promoter sequences. Some of the proteins may be useful in
CC   diagnosing and treating several disorders including, but not limited to:
CC   cancer, hyperlipidaemia, cardiovascular and neurodegenerative disorders,
CC   autoimmune diseases, and rheumatic diseases, embryogenic disorders,
CC   hypertension, renal injury, amino acidurias, hypoglycaemia, male rat
CC   infertility and myopathies.
XX
SQ   Sequence 47 BP; 10 A; 24 C; 1 G; 11 U; 1 other;


alignment_scores:
            Quality:   34.00              Length:     10
              Ratio:  3.400                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x Z40749   ..

 Align seg 1/1  to: Z40749  from: 1  to: 47

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         ||||||||||:::|||:::::::|||:::|||
       7 CUACUCCCAUCCAAUUCCACCCUAACUCCU 36

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X88170

seq_documentation_block:
ID   X88170 standard; RNA; 47 BP.
XX
AC   X88170;
XX
DT   23-SEP-1999  (first entry)
XX
DE   Oligoribonucleotide 5'm7Gppp.
XX
KW   Secreted protein; human; cytostatic; thrombotic; osteopathic; forensic;
KW   diagnostic; gene therapy; chromosome mapping; secretion vector; primer;
KW   ss.
XX
OS   Synthetic.
XX
PN   WO9925825-A2.
XX
PD   27-MAY-1999.
XX
PF   13-NOV-1998;   98WO-IB01862.
XX
PR   04-SEP-1998;   98US-0099273.
PR   13-NOV-1997;   97US-0066677.
PR   17-DEC-1997;   97US-0069957.
PR   09-FEB-1998;   98US-0074121.
PR   13-APR-1998;   98US-0081563.
PR   10-AUG-1998;   98US-0096116.
XX
PA   (GEST ) GENSET.
XX
PI   Bougueleret L,  Duclert A,  Dumas Milne Edwards J;
XX
DR   WPI; 1999-347472/29.
XX
PT   Extended cDNAs encoding secreted proteins
XX
PS   Example 2; Page 131; 307pp; English.
XX
CC   This invention describes novel nucleic acid sequences of extended cDNAs
CC   (see X97813-X97906) which encode human secreted proteins (see
CC   Y36129-Y36222) and which have cytostatic, thrombotic and osteopathic
CC   activity. The extended cDNAs can be used to express secreted proteins
CC   or parts of them or to obtain antibodies capable of binding to the
CC   secreted proteins. They may also be used in diagnostic, forensic,
CC   gene therapy and chromosome mapping procedures. Uses also include design
CC   of expression vectors and secretion vectors. This sequence represents
CC   an oligoribonucleotide primer used in the method of the invention.
XX
SQ   Sequence 47 BP; 10 A; 24 C; 2 G; 11 U; 0 other;


alignment_scores:
            Quality:   34.00              Length:     10
              Ratio:  3.400                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x X88170   ..

 Align seg 1/1  to: X88170  from: 1  to: 47

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         ||||||||||:::|||:::::::|||:::|||
       7 CUACUCCCAUCCAAUUCCACCCUAACUCCU 36

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:X97543

seq_documentation_block:
ID   X97543 standard; RNA; 47 BP.
XX
AC   X97543;
XX
DT   13-SEP-1999  (first entry)
XX
DE   Oligonucleotide Cap+ for secreted protein coding sequence isolation.
XX
KW   Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW   cellular differentiation; immune system regulator; anti-inflammatory;
KW   haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW   reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW   genetic disease; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1
FT                   /*tag= a
FT                   /mod_base= 7-methylguanosine
XX
PN   WO9931236-A2.
XX
PD   24-JUN-1999.
XX
PF   17-DEC-1998;   98WO-IB02122.
XX
PR   10-AUG-1998;   98US-0096116.
PR   17-DEC-1997;   97US-0069957.
PR   09-FEB-1998;   98US-0074121.
```

PR 13-APR-1998; 98US-0081563.
XX
PA (GEST) GENSET.
XX
PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX
DR WPI; 1999-385906/32.
XX
PT New isolated human secreted proteins
XX
PS Example 2; Page 12; 516pp; English.
XX
CC This sequence represents an oligonucleotide used to isolate the extended
CC human secreted protein coding sequences of the invention. The secreted
CC proteins can be used in treating or controlling a variety of human
CC conditions. The secreted proteins may act as cytokines or may affect
CC cellular proliferation or differentiation or may act as immune system
CC regulators, haematopoiesis regulators, tissue growth regulators,
CC regulators of reproductive hormones or cell movement or have
CC chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or tumour
CC inhibition activity. The DNAs can be used in forensic procedures to
CC identify individuals or in diagnostic procedures to identify individuals
CC having genetic diseases resulting from abnormal expression of the genes
CC corresponding to the extended cDNAs. They are also useful for
CC constructing a high resolution map of the human chromosomes. They can
CC also be used for gene therapy to control or treat genetic diseases.
XX
SQ Sequence 47 BP; 10 A; 24 C; 2 G; 11 U; 0 other;

alignment_scores:
Quality: 34.00 Length: 10
Ratio: 3.400 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 60.000

alignment_block:
US-08-860-232-1 x X97543 ..

Align seg 1/1 to: X97543 from: 1 to: 47

```
      1 LeuLeuProGluAsnAsnValLeuSerPro 10
        ||||||||||:::|||::::::|||:::|||
      7 CUACUCCCAUCCAAUUCCACCCUAACUCCU 36
```

```
OM of: US-08-860-232-1 to: Issued_Patents_NA:*   out_format : pfs

Date: Dec 12, 2000  3:35 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US08860232/runat_11122000_153406_14822/app_query.fasta_1.67
 -DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=lim60.rni
 -GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000
 -LOOPEXT=0.000 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000
 -XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000
 -YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200
 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
 -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=60
 -USER=US08860232_@CGN1_1_43 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-08-860-232-1
Query length: 11
Database: Issued_Patents_NA:*
Database sequences: 262060
Database length: 75620496
Search time (sec): 57.460000



score_list:
Sequence           Strd Orig    ZScore    EScore Len  ! Documentation  ..
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-047-041A-16 +    36.00  144.04  1.68      22  !
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-047-041A-3 +    36.00  136.10  4.65      52  !  S
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-795-006A-23 +    36.00  136.10  4.65      52  !
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-930-102A-4 +    34.00  130.12  10.02     46  !  Se
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-930-102A-5 +    34.00  130.12  10.02     46  !  Se
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-726-807B-54 +    32.00  122.80  25.61     47  !  S
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-09-258-367-54 +    32.00  122.80  25.61     47  !  Se
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-09-258-367-55 -    32.00  122.80  25.61     47  !  Se
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-171-389-8 +    31.00  119.05  41.44     48  !  Se
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-123-936-8 +    31.00  119.05  41.44     48  !  Se
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-475-228A-8 +    31.00  119.05  41.44     48  !  S
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-482-080A-8 +    31.00  119.05  41.44     48  !  Se
/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US93-12388-8 +    31.00  119.05  41.44     48  !
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-171-389-79 +    31.00  118.86  42.47     49  !  S
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-123-936-79 +    31.00  118.86  42.47     49  !  S
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-475-228A-79 +    31.00  118.86  42.47     49  !
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-482-080A-79 +    31.00  118.86  42.47     49  !  S
/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US93-12388-79 +    31.00  118.86  42.47     49
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-643-704A-10 +    31.00  116.99  53.97     60  !  S
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-388-029A-28 -    29.00  111.38  110.92    51  !  S
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-388-029A-29 -    29.00  111.38  110.92    51  !  S
/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US92-01433A-1 +    28.00  111.29  112.11    35
/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US91-00909-11 +    28.00  109.39  143.04    43
/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US91-00909-12 -    28.00  109.18  146.99    44
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-350-884-60 +    28.00  106.47  208.02    59  !  S
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-440-548-60 +    28.00  106.47  208.02    59  !  S
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-709-173-60 +    28.00  106.47  208.02    59  !  S
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-709-177-60 +    28.00  106.47  208.02    59  !  S
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-835-728D-63 +    27.00  110.13  130.14    27  !  S
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-835-728D-167 -    27.00  110.13  130.14    27  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-484-557C-23 -    27.00  109.79  135.86    28  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-487-426B-23 -    27.00  109.79  135.86    28  !
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-487-720A-23 -    27.00  109.79  135.86    28  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-334-490-2 -    27.00  109.16  147.43    30  !  Se
/cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-485-611A-2 -    27.00  109.16  147.43    30  !  S
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-07-762-136A-4 -    27.00  107.73  176.94    35  !  S
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-058-723-4 -    27.00  107.73  176.94    35  !  Se
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-411-796-362 +    27.00  107.47  182.94    36  !
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-471-039-362 +    27.00  107.47  182.94    36  !  S
/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US93-11198-362 +    27.00  107.47  182.94    36
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-653-740-34 -    27.00  106.50  207.25    40  !
/cgn2_6/ptodata/1/ina/5D_COMB.seq:US-09-073-594-34 -    27.00  106.50  207.25    40  !
/cgn2_6/ptodata/1/ina/6_COMB.seq:US-09-275-925-34 -    27.00  106.50  207.25    40  !
/cgn2_6/ptodata/1/ina/5D_COMB.seq:US-08-185-828A-10 +    27.00  104.26  276.31    51
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-225-224-16 +    27.00  103.23  315.20    57  !

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-047-041A-16

seq_documentation_block:
; Sequence 16, Application US/08047041A
; Patent No. 5527676
;  GENERAL INFORMATION:
;    APPLICANT:  Vogelstein, Bert
;    APPLICANT:  Baker, Suzanne J.
;    APPLICANT:  Fearon, Eric R.
;    APPLICANT:  Nigro, Janice M.
;    TITLE OF INVENTION:  Detection of Loss of the Wild-Type p53
;    TITLE OF INVENTION:  Gene
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Allegretti, Ltd.
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20001.4597
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/047,041A
;      FILING DATE:  22-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/928,661
;      FILING DATE:  17-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/446,584
;      FILING DATE:  06-DEC-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/330,566
;      FILING DATE:  29-MAR-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.42917
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  exon 3
;    PUBLICATION INFORMATION:
;      AUTHORS:  Lamb,
;      JOURNAL:  Mol. Cell. Biol.
;      VOLUME:  6
;      ISSUE:  5
;      PAGES:  1379-1385
;      DATE:  1986
US-08-047-041A-16
```

```
alignment_scores:
            Quality:   36.00              Length:       7
              Ratio:  5.143                Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-047-041A-16    ..

 Align seg 1/1  to: US-08-047-041A-16  from: 1  to: 22

       2 LeuProGluAsnAsnValLeu 8
         ||||||||||||||||||||
       2 CTTCCTGAAAACAACGTTCTG 22

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-047-041A-3

seq_documentation_block:
; Sequence 3, Application US/08047041A
; Patent No. 5527676
;  GENERAL INFORMATION:
;    APPLICANT:  Vogelstein, Bert
;    APPLICANT:  Baker, Suzanne J.
;    APPLICANT:  Fearon, Eric R.
;    APPLICANT:  Nigro, Janice M.
;    TITLE OF INVENTION:  Detection of Loss of the Wild-Type p53
;    TITLE OF INVENTION:  Gene
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Allegretti, Ltd.
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20001.4597
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/047,041A
;      FILING DATE:  22-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/928,661
;      FILING DATE:  17-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/446,584
;      FILING DATE:  06-DEC-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/330,566
;      FILING DATE:  29-MAR-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.42917
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  52 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  exon 3
;    PUBLICATION INFORMATION:
;      AUTHORS:  Buchman, V. L.
;      TITLE:  A variation in the structure of the
;      TITLE:  protein-coding region of the human p53 gene
;      JOURNAL:  Gene
;      VOLUME:  70
;      PAGES:  245-252
;      DATE:  1988
US-08-047-041A-3


alignment_scores:
            Quality:   36.00              Length:       7
              Ratio:  5.143                Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-047-041A-3    ..

 Align seg 1/1  to: US-08-047-041A-3  from: 1  to: 52

       2 LeuProGluAsnAsnValLeu 8
         ||||||||||||||||||||
      17 CTTCCTGAAAACAACGTTCTG 37

seq_name: /cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-795-006A-23

seq_documentation_block:
; Sequence 23, Application US/08795006A
; Patent No. 5840579
;  GENERAL INFORMATION:
;    APPLICANT:  Boeke, Jef
;    APPLICANT:  Brachmann, Rainer
;    TITLE OF INVENTION:  NUCLEIC ACIDS ENCODING P53
;    TITLE OF INVENTION:  MUTATIONS WHICH SUPPRESS P53 CANCER MUTA- TIONS
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff
;      STREET:  1001 G Street, NW
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/795,006A
;      FILING DATE:  05-FEB-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A
;      REGISTRATION NUMBER:  32141
;      REFERENCE/DOCKET NUMBER:  01107.03170
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  52 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-795-006A-23


alignment_scores:
```

```
           Quality:   36.00            Length:      7
             Ratio:  5.143              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-860-232-1 x US-08-795-006A-23   ..

 Align seg 1/1  to: US-08-795-006A-23  from: 1  to: 52

       2 LeuProGluAsnAsnValLeu 8
         ||||||||||||||||||||
      17 CTTCCTGAAAACAACGTTCTG 37

seq_name: /cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-930-102A-4

seq_documentation_block:
; Sequence 4, Application US/08930102A
; Patent No. 6022715
;  GENERAL INFORMATION:
;    APPLICANT:  Dumas, Jean-Baptiste Milne Edwards
;    APPLICANT:  Merenkova, Irena Nicolaevna
;    TITLE OF INVENTION:  METHOD FOR THE SPECIFIC COUPLING OF THE CAP
;    TITLE OF INVENTION:  OF THE 5'END OF AN mRNA FRAGMENT AND PREPARTION OF mRNA AND OF
;    TITLE OF INVENTION:  COMPLETE cDNA
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson & Bear
;      STREET:  620 Newport Center Drive, 16th Floor
;      CITY:  Newport Beach
;      STATE:  CA
;      COUNTRY:  U.S.A.
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/930,102A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR96/00651
;      FILING DATE:  29-APR-1996
;      APPLICATION NUMBER:  FR95/05221
;      FILING DATE:  02-MAY-1995
;      APPLICATION NUMBER:  FR95/09467
;      FILING DATE:  03-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ned A. Israelsen
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:  GENSET.017APC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-235-8550
;      TELEFAX:  619-235-0176
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1
;      OTHER INFORMATION:  /label= m7Gppp
;      OTHER INFORMATION:  /note= "N is m7GpppG"
US-08-930-102A-4


alignment_scores:
           Quality:   34.00            Length:     10
             Ratio:  3.400              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x US-08-930-102A-4   ..

 Align seg 1/1  to: US-08-930-102A-4  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         |||||||||:::|||:::::::|||:::|||
       6 CUACUCCCAUCCAAUUCCACCCUAACUCCU 35

seq_name: /cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-930-102A-5

seq_documentation_block:
; Sequence 5, Application US/08930102A
; Patent No. 6022715
;  GENERAL INFORMATION:
;    APPLICANT:  Dumas, Jean-Baptiste Milne Edwards
;    APPLICANT:  Merenkova, Irena Nicolaevna
;    TITLE OF INVENTION:  METHOD FOR THE SPECIFIC COUPLING OF THE CAP
;    TITLE OF INVENTION:  OF THE 5'END OF AN mRNA FRAGMENT AND PREPARTION OF mRNA AND
;    TITLE OF INVENTION:  COMPLETE cDNA
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson & Bear
;      STREET:  620 Newport Center Drive, 16th Floor
;      CITY:  Newport Beach
;      STATE:  CA
;      COUNTRY:  U.S.A.
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/930,102A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR96/00651
;      FILING DATE:  29-APR-1996
;      APPLICATION NUMBER:  FR95/05221
;      FILING DATE:  02-MAY-1995
;      APPLICATION NUMBER:  FR95/09467
;      FILING DATE:  03-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ned A. Israelsen
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:  GENSET.017APC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-235-8550
;      TELEFAX:  619-235-0176
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1
;      OTHER INFORMATION:  /label= ppp
;      OTHER INFORMATION:  /note= "N is pppG"
US-08-930-102A-5


alignment_scores:
           Quality:   34.00            Length:     10
             Ratio:  3.400              Gaps:      0
```

```
 Percent Similarity: 100.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x US-08-930-102A-5   ..

 Align seg 1/1  to: US-08-930-102A-5  from: 1  to: 46

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         ||||||||||:::|||:::::::|||:::|||
       6 CUACUCCCAUCCAAUUCCACCCUAACUCCU 35

seq_name: /cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-726-807B-54

seq_documentation_block:
; Sequence 54, Application US/08726807B
; Patent No. 6090618
;  GENERAL INFORMATION:
;    APPLICANT:  Parmacek, Michael S.
;    APPLICANT:  Solway, Julian
;    TITLE OF INVENTION:  PROMOTER FOR SMOOTH MUSCLE CELL
;    TITLE OF INVENTION:  EXPRESSION
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,807B
;      FILING DATE:  07-OCT-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/004,868
;      FILING DATE:  05-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMillian, Nabeela R.
;      REGISTRATION NUMBER:  P-43,363
;      REFERENCE/DOCKET NUMBER:  ARSB:510
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  54:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  47 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-726-807B-54


alignment_scores:
            Quality:   32.00             Length:     10
              Ratio:  3.556               Gaps:      0
 Percent Similarity:  90.000   Percent Identity: 50.000

alignment_block:
 US-08-860-232-1 x US-08-726-807B-54   ..

 Align seg 1/1  to: US-08-726-807B-54  from: 1  to: 47

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         |||:::|||   |||::::::|||:::|||
      13 CTAGTTCCCACCAACTCGATTTTAAAGCCT 42

seq_name: /cgn2_6/ptodata/1/ina/6_COMB.seq:US-09-258-367-54

seq_documentation_block:
; Sequence 54, Application US/09258367
; Patent No. 6114311
; GENERAL INFORMATION:
;  APPLICANT: PARMACEK, MICHAEL S.
;  APPLICANT: SOLWAY, JULIAN
;  TITLE OF INVENTION: PROMOTER FOR SMOOTH MUSCLE CELL EXPRESSION
;  FILE REFERENCE: ARCD:310
;  CURRENT APPLICATION NUMBER: US/09/258,367
;  CURRENT FILING DATE: 1999-02-26
;  EARLIER APPLICATION NUMBER: 08/726,807
;  EARLIER FILING DATE: 1996-10-07
;  NUMBER OF SEQ ID NOS: 55
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Primer
US-09-258-367-54


alignment_scores:
            Quality:   32.00             Length:     10
              Ratio:  3.556               Gaps:      0
 Percent Similarity:  90.000   Percent Identity: 50.000

alignment_block:
 US-08-860-232-1 x US-09-258-367-54   ..

 Align seg 1/1  to: US-09-258-367-54  from: 1  to: 47

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         |||:::|||   |||::::::|||:::|||
      13 CTAGTTCCCACCAACTCGATTTTAAAGCCT 42

seq_name: /cgn2_6/ptodata/1/ina/6_COMB.seq:US-09-258-367-55

seq_documentation_block:
; Sequence 55, Application US/09258367
; Patent No. 6114311
; GENERAL INFORMATION:
;  APPLICANT: PARMACEK, MICHAEL S.
;  APPLICANT: SOLWAY, JULIAN
;  TITLE OF INVENTION: PROMOTER FOR SMOOTH MUSCLE CELL EXPRESSION
;  FILE REFERENCE: ARCD:310
;  CURRENT APPLICATION NUMBER: US/09/258,367
;  CURRENT FILING DATE: 1999-02-26
;  EARLIER APPLICATION NUMBER: 08/726,807
;  EARLIER FILING DATE: 1996-10-07
;  NUMBER OF SEQ ID NOS: 55
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Primer
US-09-258-367-55


alignment_scores:
            Quality:   32.00             Length:     10
              Ratio:  3.556               Gaps:      0
 Percent Similarity:  90.000   Percent Identity: 50.000

alignment_block:
 US-08-860-232-1 x US-09-258-367-55/rev   ..

 Align seg 1/1  to reverse of: US-09-258-367-55  from: 1  to: 47
```

```
      1 LeuLeuProGluAsnAsnValLeuSerPro 10
        |||:::|||   |||:::::::|||:::|||
     35 CTAGTTCCCACCAACTCGATTTTAAAGCCT 6

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-171-389-8

seq_documentation_block:
; Sequence 8, Application US/08171389
; Patent No. 5578444
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,389
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175/G19P3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human choline acetyltransferase gene
US-08-171-389-8


alignment_scores:
            Quality:   31.00            Length:      9
              Ratio:  4.429              Gaps:      0
 Percent Similarity:  77.778   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x US-08-171-389-8   ..

 Align seg 1/1  to: US-08-171-389-8  from: 1  to: 48

       3 ProGluAsnAsnValLeuSerProLeu 11
         |||:::||||||   |||   ||||||
      18 CCTAATAATAACAGTCTTTGCCCTCTT 44

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-123-936-8

seq_documentation_block:
; Sequence 8, Application US/08123936
; Patent No. 5726014
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    TITLE OF INVENTION:  Screening Assay for the Detection of
;    TITLE OF INVENTION:  DNA-Binding Molecules
;    NUMBER OF SEQUENCES:  640
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/123,936
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0075.32/G19P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human choline acetyltransferase gene
US-08-123-936-8


alignment_scores:
            Quality:   31.00            Length:      9
              Ratio:  4.429              Gaps:      0
 Percent Similarity:  77.778   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x US-08-123-936-8   ..

 Align seg 1/1  to: US-08-123-936-8  from: 1  to: 48
```

```
      3 ProGluAsnAsnValLeuSerProLeu 11
        |||:::||||||   |||   ||||||
     18 CCTAATAATAACAGTCTTTGCCCTCTT 44

seq_name: /cgn2_6/ptodata/1/ina/5C_COMB.seq:US-08-475-228A-8

seq_documentation_block:
; Sequence 8, Application US/08475228A
; Patent No. 5869241
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,228A
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stratford, Carol A.
;      REGISTRATION NUMBER:  34,444
;      REFERENCE/DOCKET NUMBER:  4600-0175.21/G19P3D2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human choline acetyltransferase gene
US-08-475-228A-8


alignment_scores:
           Quality:   31.00              Length:      9
             Ratio:  4.429                Gaps:       0
 Percent Similarity:  77.778   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x US-08-475-228A-8    ..
```

```
 Align seg 1/1  to: US-08-475-228A-8  from: 1  to: 48

      3 ProGluAsnAsnValLeuSerProLeu 11
        |||:::||||||   |||   ||||||
     18 CCTAATAATAACAGTCTTTGCCCTCTT 44

seq_name: /cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-482-080A-8

seq_documentation_block:
; Sequence 8, Application US/08482080A
; Patent No. 6010849
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,080A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brady, John F.
;      REGISTRATION NUMBER:  39,118
;      REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-0880
;      TELEFAX:  (650) 324-0960
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human choline acetyltransferase gene
US-08-482-080A-8


alignment_scores:
           Quality:   31.00              Length:      9
             Ratio:  4.429                Gaps:       0
```

```
 Percent Similarity:  77.778   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x US-08-482-080A-8   ..

 Align seg 1/1  to: US-08-482-080A-8  from: 1  to: 48

       3 ProGluAsnAsnValLeuSerProLeu 11
         |||:::||||||   |||   ||||||
      18 CCTAATAATAACAGTCTTTGCCCTCTT 44

seq_name: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US93-12388-8

seq_documentation_block:
; Sequence 8, Application PC/TUS9312388
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12388
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175.41/G19PCT2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human choline acetyltransferase gene
PCT-US93-12388-8


alignment_scores:
            Quality:   31.00             Length:      9
              Ratio:  4.429               Gaps:      0
 Percent Similarity:  77.778   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x PCT-US93-12388-8   ..

 Align seg 1/1  to: PCT-US93-12388-8  from: 1  to: 48

       3 ProGluAsnAsnValLeuSerProLeu 11
         |||:::||||||   |||   ||||||
      18 CCTAATAATAACAGTCTTTGCCCTCTT 44

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-171-389-79

seq_documentation_block:
; Sequence 79, Application US/08171389
; Patent No. 5578444
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,389
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175/G19P3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  79:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human choline acetyltransferase gene
US-08-171-389-79


alignment_scores:
            Quality:   31.00             Length:      9
              Ratio:  4.429               Gaps:      0
 Percent Similarity:  77.778   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x US-08-171-389-79   ..
```

```
 Align seg 1/1  to: US-08-171-389-79  from: 1  to: 49

       3 ProGluAsnAsnValLeuSerProLeu 11
         |||:::||||||   |||   ||||||
      19 CCTAATAATAACAGTCTTTGCCCTCTT 45

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-123-936-79

seq_documentation_block:
; Sequence 79, Application US/08123936
; Patent No. 5726014
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    TITLE OF INVENTION:  Screening Assay for the Detection of
;    TITLE OF INVENTION:  DNA-Binding Molecules
;    NUMBER OF SEQUENCES:  640
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/123,936
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0075.32/G19P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  79:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human choline acetyltransferase gene
US-08-123-936-79


alignment_scores:
            Quality:   31.00              Length:      9
              Ratio:  4.429                Gaps:      0
 Percent Similarity:  77.778   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x US-08-123-936-79   ..

 Align seg 1/1  to: US-08-123-936-79  from: 1  to: 49

       3 ProGluAsnAsnValLeuSerProLeu 11
         |||:::||||||   |||   ||||||
      19 CCTAATAATAACAGTCTTTGCCCTCTT 45
```

```
OM of: US-08-860-232-1 to: EST:*   out_format : pfs

Date: Dec 12, 2000  3:19 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US08860232/runat_11122000_153405_14796/app_query.fasta_1.67
 -DB=EST -QFMT=fastap -SUFFIX=lim60.rst -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=60 -USER=US08860232_@CGN1_1_1780
 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-08-860-232-1
Query length: 11
Database: EST:*
Database sequences: 7189864
Database length: -1203564053
Search time (sec): 415.240000


score_list:
Sequence              Strd Orig    ZScore    EScore Len   ! Documentation  ..
gb_est8:AI156636        +   30.00   116.19  1.7e+03   34   ! AI156636 ue53a12.r1 Soares_mamma
gb_est9:AI290158        +   29.00   112.65  2.6e+03   34   ! AI290158 qw21c03.x1 NCI_CGAP_Ut4
gb_est5:AA587946        -   28.00   106.41  5.8e+03   46   ! AA587946 nm88f04.s1 NCI_CGAP_Co9
gb_est2:AA196079        +   28.00   104.98  7.0e+03   54   ! AA196079 zp98d07.s1 Stratagene m
gb_est8:AI150170        -   28.00   104.82  7.1e+03   55   ! AI150170 qf34b07.x1 Soares_testi
gb_est11:AI614604       +   28.00   104.66  7.3e+03   56   ! AI614604 mm33f01.y1 Stratagene m
gb_est5:AA641043        -   28.00   104.34  7.6e+03   58   ! AA641043 nr25a10.r1 NCI_CGAP_Pr2
gb_est7:AA878695        -   27.00   103.46  8.5e+03   43   ! AA878695 oj23c05.s1 NCI_CGAP_Kid
gb_est5:AA683896        +   27.00   102.12  1.0e+04   50   ! AA683896 vr06e08.r1 Knowles Solt
gb_est37:H45714         -   27.00   101.27  1.1e+04   55   ! H45714 yp23a05.r1 Soares breast
gb_est37:F31384         -   27.00   100.80  1.2e+04   58   ! F31384 HSPD22527 HM3 Homo sapien
gb_gss19:B02861         +   27.00   100.80  1.2e+04   58   ! B02861 cSRL-161D12-u cSRL flow s
gb_est2:AA206512        +   26.00   100.56  1.2e+04   40   ! AA206512 zq56f04.r1 Stratagene m
gb_est11:AI538057       -   26.00    99.51  1.4e+04   45   ! AI538057 to83b07.x1 NCI_CGAP_Gas
gb_est6:AA854290        +   26.00    99.32  1.4e+04   46   ! AA854290 aj65d05.s1 Soares_parat
gb_est12:AI748865       +   26.00    99.32  1.4e+04   46   ! AI748865 at37b11.x1 Barstead col
gb_est5:AA584618        +   26.00    97.89  1.7e+04   54   ! AA584618 no08h03.s1 NCI_CGAP_Phe
gb_est7:AA910226        -   26.00    97.72  1.8e+04   55   ! AA910226 oj29f04.s1 NCI_CGAP_Kid
gb_est21:AW248622       +   26.00    97.56  1.8e+04   56   ! AW248622 2820745.3prime NIH_MGC_
gb_est38:R42058         -   26.00    97.25  1.9e+04   58   ! R42058 yg05h08.s1 Soares infant
gb_est5:AA593486        -   26.00    96.95  2.0e+04   60   ! AA593486 nn26h11.s1 NCI_CGAP_Gas
gb_gss19:B03225         +   26.00    96.95  2.0e+04   60   ! B03225 cSRL-173B8-u cSRL flow sc
gb_est7:AA991999        +   25.00    98.46  1.6e+04   34   ! AA991999 ot01f02.s1 NCI_CGAP_GC3
gb_est7:AA932036        +   25.00    97.01  1.9e+04   40   ! AA932036 oo35g08.s1 NCI_CGAP_Lu5
gb_est7:AA947665        -   25.00    96.37  2.1e+04   43   ! AA947665 oq32d02.s1 NCI_CGAP_GC4
gb_est12:AI685223       -   25.00    96.37  2.1e+04   43   ! AI685223 wc70e01.x1 NCI_CGAP_Pan
gb_est14:AJ237252       +   25.00    95.21  2.4e+04   49   ! AJ237252 AJ237252 Emericella nid
gb_est35:BE387699       +   25.00    95.03  2.5e+04   50   ! BE387699 601287355F1 NIH_MGC_44
gb_est6:AA846708        +   25.00    94.51  2.7e+04   53   ! AA846708 aj41b12.s1 Soares_testi
gb_est4:AA515004        +   25.00    94.18  2.8e+04   55   ! AA515004 nf54e03.s1 NCI_CGAP_Alv
gb_est7:AA878824        -   25.00    94.18  2.8e+04   55   ! AA878824 of87c08.s1 NCI_CGAP_Li5
gb_est7:AA972473        -   25.00    94.18  2.8e+04   55   ! AA972473 op42a11.s1 Soares_NFL_T
gb_est9:AI195165        +   25.00    94.18  2.8e+04   55   ! AI195165 ui61d09.x1 Sugano mouse
gb_est12:AI664266       -   25.00    94.18  2.8e+04   55   ! AI664266 ue83e10.r1 Soares_NMPu
gb_gss19:B05394         -   25.00    94.18  2.8e+04   55   ! B05394 cSRL-61e3-u cSRL flow sor
gb_gss24:GGA350127      +   25.00    94.18  2.8e+04   55   ! AJ232133 Gallus gallus anonymous
gb_est19:AV554548       +   25.00    94.02  2.9e+04   56   ! AV554548 AV554548 Arabidopsis th
gb_est1:AA119098        -   25.00    93.71  3.0e+04   58   ! AA119098 mp62c05.r1 Soares_thymu
gb_est4:AA412333        +   25.00    93.71  3.0e+04   58   ! AA412333 zt97h02.r1 Soares_testi
gb_est10:AI441338       -   25.00    93.71  3.0e+04   58   ! AI441338 sa55e03.y1 Gm-c1004 Gly
gb_est24:AW708404       +   25.00    93.55  3.0e+04   59   ! AW708404 c6d11ne.f1 Neurospora c
gb_est36:D19151         +   24.00    97.31  1.9e+04   26   ! D19151 MUSGS01373 Mouse 3'-di
gb_est7:AA907811        +   24.00    96.65  2.0e+04   28   ! AA907811 og01g04.s2 NCI_CGAP_
gb_est11:AI500674       +   24.00    94.16  2.8e+04   37   ! AI500674 tn99b01.x1 NCI_CGAP_
gb_est39:T52826         +   24.00    94.16  2.8e+04   37   ! T52826 ya81b02.s1 Stratagene

seq_name: gb_est8:AI156636

seq_documentation_block:
LOCUS       AI156636       34 bp    mRNA            EST       30-SEP-1998
DEFINITION  ue53a12.r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
            IMAGE:1494814 5' similar to SW:NIMM_BOVIN Q02377 NADH-UBIQUINONE
            OXIDOREDUCTASE MFWE SUBUNIT ;, mRNA sequence.
ACCESSION   AI156636
VERSION     AI156636.1  GI:3685105
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:932418
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .34
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1494814"
                     /clone_lib="Soares_mammary_gland_NMLMG"
                     /sex="female (lactating)"
                     /tissue_type="mammary gland"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from mammary
                     gland tissue from a lactating female, and was then primed
                     with a Not I - oligo(dT) primer. Double-stranded  cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I  and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized.  Library
                     was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT        8 a      7 c      9 g     10 t
ORIGIN


alignment_scores:
            Quality:   30.00            Length:      8
              Ratio:  3.750              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 62.500

alignment_block:
 US-08-860-232-1 x AI156636   ..

 Align seg 1/1  to: AI156636  from: 1  to: 34

       1 LeuLeuProGluAsnAsnValLeu 8
```

```
          :::||||||||||||::::::|||
      11 ATTCTCCCTGAGAATAGCATGCTA 34

seq_name: gb_est9:AI290158

seq_documentation_block:
LOCUS       AI290158       34 bp    mRNA            EST       30-NOV-1998
DEFINITION  qw21c03.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:1991716 3'
            similar to TR:Q15214 Q15214 SALIVARY PROLINE-RICH PROTEIN 1 ;, mRNA
            sequence.
ACCESSION   AI290158
VERSION     AI290158.1  GI:3931824
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .34
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1991716"
                     /clone_lib="NCI_CGAP_Ut4"
                     /tissue_type="serous papillary carcinoma, high grade, 2
                     pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.48 kb. Life Technologies catalog #:
                     11542-016"
BASE COUNT        7 a     19 c      2 g      6 t
ORIGIN


alignment_scores:
           Quality:   29.00             Length:     10
             Ratio:  3.625               Gaps:      0
 Percent Similarity:  80.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x AI290158   ..

 Align seg 1/1  to: AI290158  from: 1  to: 34

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         |||   |||:::::|||   |||||||||
       2 CTCTGCCCCAACAAAAACCCCCTTTCCCCC 31

seq_name: gb_est5:AA587946

seq_documentation_block:
LOCUS       AA587946       46 bp    mRNA            EST       25-SEP-1997
DEFINITION  nm88f04.s1 NCI_CGAP_Co9 Homo sapiens cDNA clone IMAGE:1075327 3'
            similar to gb:X59268 TRANSCRIPTION INITIATION FACTOR IIB (HUMAN);,
            mRNA sequence.
ACCESSION   AA587946
VERSION     AA587946.1  GI:2402121
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.
            , Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 616   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1075327"
                     /clone_lib="NCI_CGAP_Co9"
                     /tissue_type="colon tumor RER+"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     RER+ colon tumor, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is not normalized. Library was
                     constructed by Bento Soares and M. Fatima Bonaldo (Soares4
                     ). "
BASE COUNT       15 a      2 c     19 g     10 t
ORIGIN


alignment_scores:
           Quality:   28.00             Length:     10
             Ratio:  3.500               Gaps:      0
 Percent Similarity:  80.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x AA587946/rev   ..

 Align seg 1/1  to reverse of: AA587946  from: 1  to: 46

       2 LeuProGluAsnAsnValLeuSerProLeu 11
         |||||||:::   :::   ||||||||||||
      41 TTACCACGCTACTCCTACCTATCTCCCCTT 12

seq_name: gb_est2:AA196079

seq_documentation_block:
LOCUS       AA196079       54 bp    mRNA            EST       10-MAR-1998
DEFINITION  zp98d07.s1 Stratagene muscle 937209 Homo sapiens cDNA clone
            IMAGE:628237 3' similar to gb:M73628 KAPPA CASEIN PRECURSOR (HUMAN
            );, mRNA sequence.
ACCESSION   AA196079
```

```
VERSION     AA196079.1  GI:1791662
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 797   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Homo sapiens"
                     /db_xref="GDB:5048471"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:628237"
                     /clone_lib="Stratagene muscle 937209"
                     /tissue_type="muscle"
                     /dev_stage="adult"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skeletal muscle; Vector: pBluescript SK-;
                     Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally.
                     Primer: Oligo dT. Skeletal muscle from patient with
                     malignant hyperthermia.  Average insert size: 1.0 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a     12 c      7 g     13 t      3 others
ORIGIN


alignment_scores:
            Quality:   28.00             Length:     10
              Ratio:  4.000               Gaps:      0
 Percent Similarity:  70.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x AA196079   ..

 Align seg 1/1  to: AA196079  from: 1  to: 54

       1 LeuLeuProGluAsnAsnValLeuSerPro 10
         :::|||      ||||||   |||||||||
      22 ATTCTNTCTTGGAACAACCNGCTATCACCA 51

seq_name: gb_est8:AI150170

seq_documentation_block:
LOCUS       AI150170       55 bp    mRNA            EST       10-NOV-1998
DEFINITION  qf34b07.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1751893
            3' similar to gb:X59268 TRANSCRIPTION INITIATION FACTOR IIB (HUMAN
            );contains TAR1.b1 MSR1 repetitive element ;, mRNA sequence.
ACCESSION   AI150170
VERSION     AI150170.1  GI:3678639
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
            , Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 290   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 21.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1751893"
                     /clone_lib="Soares_testis_NHT"
                     /sex="male"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech Laboratories
                     , Inc., and primed with  a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3']. 
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       17 a      4 c     21 g     13 t
ORIGIN


alignment_scores:
            Quality:   28.00             Length:     10
              Ratio:  3.500               Gaps:      0
 Percent Similarity:  80.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x AI150170/rev   ..

 Align seg 1/1  to reverse of: AI150170  from: 1  to: 55

       2 LeuProGluAsnAsnValLeuSerProLeu 11
         ||||||:::   :::   ||||||||||||
      41 TTACCACGCTACTCCTACCTATCTCCCCTT 12

seq_name: gb_est11:AI614604

seq_documentation_block:
LOCUS       AI614604       56 bp    mRNA            EST       15-MAR-2000
DEFINITION  mm33f01.y1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
            IMAGE:523321 5', mRNA sequence.
ACCESSION   AI614604
VERSION     AI614604.1  GI:4623771
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
```

```
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            This read is a RESEQUENCE of a previously sequenced mouse clone
            This read has been verified (found to hit its original self in the
            correct orientation)
            Putative full length read
            vector to vector length is
            MGI:317169
            Seq primer: -40RP from Gibco
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:523321"
                     /clone_lib="Stratagene mouse skin (#937313)"
                     /sex="females"
                     /tissue_type="whole skin"
                     /dev_stage="11 weeks old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. Whole skin from 11 week old C57BL/6 female mice.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       18 a     15 c      4 g     19 t
ORIGIN


alignment_scores:
           Quality:   28.00            Length:     11
             Ratio:  3.111              Gaps:      0
 Percent Similarity:  81.818   Percent Identity: 63.636

alignment_block:
 US-08-860-232-1 x AI614604    ..

 Align seg 1/1  to: AI614604  from: 1  to: 56

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         ||||||||||:::   ||||||||||:::   |||
       3 CTCCTACCATCTCCAAATGTCCTTAAATGTTTA 35

seq_name: gb_est5:AA641043

seq_documentation_block:
LOCUS       AA641043       58 bp    mRNA            EST       27-OCT-1997
DEFINITION  nr25a10.r1 NCI_CGAP_Pr2 Homo sapiens cDNA clone IMAGE:1168986
            similar to gb:U02389 MITOCHONDRIAL LON PROTEASE HOMOLOG PRECURSOR
            (HUMAN);, mRNA sequence.
ACCESSION   AA641043
VERSION     AA641043.1  GI:2566293
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.
            , Michael Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1168986"
                     /clone_lib="NCI_CGAP_Pr2"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated, total cellular RNA obtained from 5,000-10
                     ,000 microdissected preneoplastic cells
                     histologically-determined to be prostatic intraepithelial
                     neoplasia 2 (PIN2) cells.  Double-stranded cDNA was
                     ligated to EcoRI adaptors, 5 cycles of PCR applied to the
                     cDNA with an adaptor-specific primer, and the resulting
                     PCR product subcloned into pAMP10 by the UDG-cloning
                     method (Life Technologies).  Average insert size is 600
                     bp. NOTE: Not directionally cloned.  This library was
                     constructed by  David Krizman."
BASE COUNT       10 a      2 c     28 g     18 t
ORIGIN


alignment_scores:
           Quality:   28.00            Length:      8
             Ratio:  4.000              Gaps:      0
 Percent Similarity:  87.500   Percent Identity: 62.500

alignment_block:
 US-08-860-232-1 x AA641043/rev    ..

 Align seg 1/1  to reverse of: AA641043  from: 1  to: 58

       3 ProGluAsnAsnValLeuSerPro 10
         |||   :::|||:::||||||||||
      50 CCACCCCATAATACTCTTTCACCC 27

seq_name: gb_est7:AA878695

seq_documentation_block:
LOCUS       AA878695       43 bp    mRNA            EST       19-MAY-1998
DEFINITION  oj23c05.s1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1493000 3'
            similar to TR:Q14626 Q14626 INCOMPLETE INTERLEUKIN-11 RECEPTOR
            ISOFORM ;, mRNA sequence.
ACCESSION   AA878695
VERSION     AA878695.1  GI:2987660
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
```

```
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 1289   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .43
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1493000"
                      /clone_lib="NCI_CGAP_Kid5"
                      /tissue_type="2 pooled tumors (clear cell type)"
                      /lab_host="DH10B"
                      /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                      a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      AACTGGAAGAATTCGCGGCCGCAATATTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT        8 a      9 c     17 g      9 t
ORIGIN


alignment_scores:
            Quality:   27.00             Length:      9
              Ratio:  3.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 55.556

alignment_block:
 US-08-860-232-1 x AA878695/rev   ..

 Align seg 1/1  to reverse of: AA878695  from: 1  to: 43

       2 LeuProGluAsnAsnValLeuSerPro 10
         |||||||||:::::::::|||:::|||
      41 CTTCCTGAAACCAGGACACTGGGTCCC 15

seq_name: gb_est5:AA683896

seq_documentation_block:
LOCUS        AA683896        50 bp    mRNA            EST       09-DEC-1997
DEFINITION   vr06e08.r1 Knowles Solter mouse blastocyst B3 Mus musculus cDNA
             clone IMAGE:1111046 5' similar to gb:X14678 Mouse TPA-induced TIS11
             mRNA (MOUSE);, mRNA sequence.
ACCESSION    AA683896
VERSION      AA683896.1  GI:2670482
KEYWORDS     EST.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 50)
  AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
             Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
             Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
             Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
             Waterston,R.
  TITLE      The WashU-HHMI Mouse EST Project
  JOURNAL    Unpublished (1996)
COMMENT      Contact: Marra M/Mouse EST Project
             WashU-HHMI Mouse EST Project
             Washington University School of MedicineP
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:609214
             Trace considered overall poor quality
             High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .50
                      /organism="Mus musculus"
                      /strain="B6D2 F1/J"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:1111046"
                      /clone_lib="Knowles Solter mouse blastocyst B3"
                      /tissue_type="blastocyst"
                      /dev_stage="embryo (pre-implantation)"
                      /lab_host="DH10B"
                      /note="Organ: embryo; Vector: pSPORT; Site_1: NotI;
                      Site_2: SalI; Cloned unidirectionally from mRNA prepared
                      from 800 blastocysts.   Primer: SalI(dT):
                      5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTT-3'.  cDNAs  were
                      cloned into the NotI/SalI sites of a pSPORT vector (Life
                      Technologies).  Two different size selections:  B1 (larger
                      inserts) and B3."
BASE COUNT       21 a      5 c     10 g     14 t
ORIGIN


alignment_scores:
            Quality:   27.00             Length:      9
              Ratio:  3.375               Gaps:      0
 Percent Similarity:  88.889   Percent Identity: 66.667

alignment_block:
 US-08-860-232-1 x AA683896   ..

 Align seg 1/1  to: AA683896  from: 1  to: 50

       1 LeuLeuProGluAsnAsnValLeuSer 9
         ||||||   ::::::||||||||||||
      18 TTACTGAGTAAGAGCAATGTGCTTTCC 44

seq_name: gb_est37:H45714

seq_documentation_block:
LOCUS        H45714          55 bp    mRNA            EST       31-JUL-1995
DEFINITION   yp23a05.r1 Soares breast 3NbHBst Homo sapiens cDNA clone
             IMAGE:188240 5' similar to SP:NIIM_BOVIN Q02375 NADH-UBIQUINONE
             OXIDOREDUCTASE 18 KD SUBUNIT PRECURSOR ;, mRNA sequence.
ACCESSION    H45714
VERSION      H45714.1  GI:921766
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 55)
  AUTHORS    Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
             ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
             Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
             ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE      The WashU-Merck EST Project
  JOURNAL    Unpublished (1995)
COMMENT      Contact: Wilson RK
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
```

```
            Insert Size: 575
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 575    Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .55
                      /organism="Homo sapiens"
                      /db_xref="GDB:3819137"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:188240"
                      /clone_lib="Soares breast 3NbHBst"
                      /sex="Female"
                      /dev_stage="adult"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: breast; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of a modified pT7T3 vector (Pharmacia).
                      Library went through one round of normalization to a Cot =
                      20. Library constructed by Bento Soares and M.Fatima
                      Bonaldo."
BASE COUNT       19 a      9 c     13 g     12 t      2 others
ORIGIN


alignment_scores:
            Quality:   27.00              Length:      9
              Ratio:  3.857                Gaps:       0
 Percent Similarity:  77.778   Percent Identity: 55.556

alignment_block:
 US-08-860-232-1 x H45714/rev   ..

 Align seg 1/1  to reverse of: H45714  from: 1  to: 55

       2 LeuProGluAsnAsnValLeuSerPro 10
         ||||||      ||||||::::::|||
      28 CTTCCTCTCNTCAATGTCATAGCTCCA 2

seq_name: gb_est37:F31384

seq_documentation_block:
LOCUS       F31384          58 bp    mRNA            EST       13-MAY-1999
DEFINITION  HSPD22527 HM3 Homo sapiens cDNA clone s4000115C12, mRNA sequence.
ACCESSION   F31384
VERSION     F31384.1  GI:4817010
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Lanfranchi,G., Muraro,T., Caldara,F., Pacchioni,B., Pallavicini,A.,
            Pandolfo,D., Toppo,S., Trevisan,S., Scarso,S. and Valle,G.
  TITLE     Identification of 4370 expressed sequence tags from a
            3'-end-specific cDNA library of human skeletal muscle by DNA
            sequencing and filter hybridization
  JOURNAL   Genome Res. 6 (1), 35-42 (1996)
  MEDLINE   96276048
COMMENT     Contact: Valle G.
            CRIBI Biotechnology Centre
            University of Padua
            Via Trieste 75, 35121 Padua, Italy
            ABI Chromatograms and other information are available on WWW at
            http://grup.bio.unipd.it.
FEATURES              Location/Qualifiers
     source           1. .58
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="s4000115C12"
                      /clone_lib="HM3"
                      /sex="female"
                      /tissue_type="pectoral muscle (after mastectomy)"
                      /note="Vector: pcDNAII (Invitrogen); Site_1: BstXI;
                      Site_2: NotI; The library was constructed by G.
                      Lanfranchi. This library is not subtracted nor normalized.
                      The first strand cDNA was primed with a biotinylated
                      oligo-dT-NotI primer
                      (5'-biotin-AACCCGGCTCGAGCGGCCGCTTTTTTTTTTTTTTTT-3'). The
                      ds cDNA was sonicated and size-selected in the range
                      350-550  bp. The 3' specific fragments were selected by
                      streptavidin coated magnetic beads, ligated to
                      non-palindromic BstXI adapters, NotI digested and
                      directionally cloned into BstXI-NotI cut pcDNAII vector."
BASE COUNT       20 a      7 c     16 g     15 t
ORIGIN


alignment_scores:
            Quality:   27.00              Length:      11
              Ratio:  2.700                Gaps:       0
 Percent Similarity:  90.909   Percent Identity: 36.364

alignment_block:
 US-08-860-232-1 x F31384/rev   ..

 Align seg 1/1  to reverse of: F31384  from: 1  to: 58

       1 LeuLeuProGluAsnAsnValLeuSerProLeu 11
         :::::::|||   :::|||::::::::::|||||
      37 GTAATACCATTTTCGAATACCATCCAGCCACTT 5

seq_name: gb_gss19:B02861

seq_documentation_block:
LOCUS       B02861          58 bp    DNA             GSS       13-JUL-1996
DEFINITION  cSRL-161D12-u cSRL flow sorted Chromosome 11 specific cosmid Homo
            sapiens genomic clone cSRL-161D12, DNA sequence.
ACCESSION   B02861
VERSION     B02861.1  GI:1412139
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Evans,G.A., Burbee,D., Davies,C., Hahner,L., Oliver,T., Gilbert,M.,
            Jones,D., Ward,T., Gillilan,E., Schagemann,J., Probst,S. ., Harris
            ,J., DeFord,J., McFarland,J., Burzinski,K., Khan,M., Kupfer,K. and
            Garner,H.R.
  TITLE     Genomic Sequence Sampled Map of Chromosome 11
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Evans GA, Shane Probst
            McDermott Center for Human Growth and Development
            University of Texas Southwestern Medical Center At Dallas
            5323 Harry Hines Blvd, Dallas TX 75235-8591
            Tel: 214-648-1600
            Fax: 214-648-1666
            Email: gevans@utsw.swmed.edu, shane@mcdermott.swmed.edu
            Seq primer: T7
            Class: cosmid ends
            High quality sequence stop: 58.
FEATURES              Location/Qualifiers
     source           1. .58
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
```

```
                     /clone="cSRL-161D12"
                     /clone_lib="cSRL flow sorted Chromosome 11 specific
                     cosmid"
                     /sex="female"
                     /cell_type="chimeric hamster somatic cell hybrid"
                     /note="Vector: sCos-1; Human Chromosome 11 specific cosmid
                     library prepared from flow sorted human  Chromosome 11
                     derived from Chinese Hampster Ovary (CHO) monochromosomal
                     somatic cell hybrid, J1"
BASE COUNT       17 a     18 c      9 g     13 t      1 others
ORIGIN


alignment_scores:
            Quality:   27.00             Length:      9
              Ratio:  3.857               Gaps:      0
 Percent Similarity:  77.778   Percent Identity: 44.444

alignment_block:
 US-08-860-232-1 x B02861   ..

 Align seg 1/1  to: B02861  from: 1  to: 58

       2 LeuProGluAsnAsnValLeuSerPro 10
         ||||||:::   |||::::::   |||
       9 CTGCCTAAAGCCAATATCATACCTCCC 35

seq_name: gb_est2:AA206512

seq_documentation_block:
LOCUS       AA206512       40 bp    mRNA            EST       27-JAN-1997
DEFINITION  zq56f04.r1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA
            clone IMAGE:645631 5' similar to TR:G1345404 G1345404 TFIID SUBUNIT
            P22. ;, mRNA sequence.
ACCESSION   AA206512
VERSION     AA206512.1  GI:1801893
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Homo sapiens"
                     /db_xref="GDB:5215882"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:645631"
                     /clone_lib="Stratagene neuroepithelium (#937231)"
                     /dev_stage="Ntera-2/RA neuroepithelial cells"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. NT2
                     cells (Ntera-2/c1.D1) induced with Retinoic Acid for 24
                     hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       10 a     10 c     12 g      8 t
ORIGIN


alignment_scores:
            Quality:   26.00             Length:      6
              Ratio:  5.200               Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 83.333

alignment_block:
 US-08-860-232-1 x AA206512   ..

 Align seg 1/1  to: AA206512  from: 1  to: 40

       1 LeuLeuProGluAsnAsn 6
         |||   ||||||||||||
      20 CTTAGCCCTGAAAACAAT 37

seq_name: gb_est11:AI538057

seq_documentation_block:
LOCUS       AI538057       45 bp    mRNA            EST       13-MAY-1999
DEFINITION  to83b07.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2184853 3'
            similar to gb:X59268 TRANSCRIPTION INITIATION FACTOR IIB (HUMAN
            );contains element MSR1 repetitive element ;, mRNA sequence.
ACCESSION   AI538057
VERSION     AI538057.1  GI:4452192
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 1403   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2184853"
                     /clone_lib="NCI_CGAP_Gas4"
                     /tissue_type="poorly differentiated adenocarcinoma with
                     signet ring cell features"
                     /lab_host="DH10B"
                     /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.69 kb. Life Technologies catalog #:
                     11549-011"
BASE COUNT       15 a      4 c     16 g     10 t
```

```
ORIGIN


alignment_scores:
            Quality:   26.00              Length:     10
              Ratio:  3.714                Gaps:      0
 Percent Similarity:  70.000   Percent Identity: 60.000

alignment_block:
 US-08-860-232-1 x AI538057/rev   ..

 Align seg 1/1  to reverse of: AI538057  from: 1  to: 45

       2 LeuProGluAsnAsnValLeuSerProLeu 11
         ||||||      :::   ||||||||||||
      41 TTACCAGGGTACTCCTACCTATCTCCCCTT 12

seq_name: gb_est6:AA854290

seq_documentation_block:
LOCUS       AA854290       46 bp    mRNA            EST       31-DEC-1998
DEFINITION  aj65d05.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
            IMAGE:1401321 3' similar to SW:S111_HUMAN P31949 CALGIZZARIN ;,
            mRNA sequence.
ACCESSION   AA854290
VERSION     AA854290.1  GI:2941828
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
            , Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 1438   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1401321"
                     /clone_lib="Soares_parathyroid_tumor_NbHPA"
                     /tissue_type="parathyroid tumor"
                     /dev_stage="adult"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
                     ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCACCAATTTTTTTTTTTTTTTTTT
                     TTTTT-3'], double-stranded cDNA was size selected, ligated
                     to Eco RI adapters (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of a modified pT7T3
                     vector (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
                     adenomas was kindly provided by Dr. Stephen Marx, National
                     Institute of Diabetes and Digestive and Kidney Diseases,
                     NIH."
BASE COUNT       12 a      13 c       7 g      14 t
ORIGIN


alignment_scores:
            Quality:   26.00              Length:     10
              Ratio:  2.889                Gaps:      0
 Percent Similarity:  90.000   Percent Identity: 40.000

alignment_block:
 US-08-860-232-1 x AA854290   ..

 Align seg 1/1  to: AA854290  from: 1  to: 46

       2 LeuProGluAsnAsnValLeuSerProLeu 11
         ||||||::::::::::   |||:::|||::::
      10 CTACCTGACCATCACTGCTTGGGTCCAGTT 39
```